



## Informal formative assessment: The role of instructional dialogues in assessing students' learning

Maria Araceli Ruiz-Primo

Laboratory of Educational Assessment, Research, and Innovation (LEARN), University of Colorado Denver, Denver, CO, United States

### ARTICLE INFO

#### Keywords:

Informal formative assessment  
Assessment conversations  
Instructional dialogues

### ABSTRACT

This paper focuses on an unceremonious type of formative assessment – *informal formative assessment* – in which much of what teachers and students do in the classroom can be described as potential assessments that can provide evidence about the students' level of understanding. More specifically, the paper focuses on **assessment conversations, or dialogic interactions or exchanges, which continuously happen in the classroom** and that are at the center of informal formative assessment. It is argued that assessment conversations make students' thinking explicit in an unobtrusive manner, and when students' thinking is explicit, it can be examined, questioned, and shaped as an active object of constructive learning. The paper conceptualizes informal formative assessment at the center of effective instructional activities with the use of instructional dialogues as assessment conversations, a typical informal formative assessment practice. The paper then presents a discussion about the evidence on the effect of assessment conversations on student learning.

© 2011 Elsevier Ltd. All rights reserved.

The premise of this paper is that formative assessment, *assessment for learning*, is a critical component of effective instructional practices (Black & Wiliam, 1998; Daws & Singh, 1996; Gearhart et al., 2006; Pellegrino, Chudowsky, & Glaser, 2001). In order to advance their students' learning, it is critical for teachers to identify potential difficulties that hamper the achievement of their learning goals, using any combination of informal and formal strategies (AFL, 2009). When teachers implement instructional practices effectively, formative assessment occurs informally in real time every day, as well as formally at pre-specified times (AFL, 2009; Leahy, Lyon, Thompson, & Wiliam, 2005).

This paper focuses on the unceremonious informal formative assessment practices that teachers employ every day. Attention to informal formative assessment practices may help teacher education programs, school districts, and state education departments to focus on formative assessment from different angles rather than focusing *only* on benchmark/on-demand/interim assessments. The following four ideas govern the body of this paper: (1) *how* informal formative assessment can be conceptualized, (2) *how* instructional dialogues can be considered informal formative assessment sources of information, (3) *what* evidence exists regarding the effect of informal formative assessment on student learning, and (4) *where* we are in our knowledge about informal formative assessment and what we can do to advance it.

### Conceptualizing informal formative assessment

Informal formative assessment is based on the idea that much of what teachers and students do in their classrooms can be described as potential assessment opportunities for collecting evidence of students' understanding (Assessment Reform Group, 2002; Black, 2009; Jordan & Putz, 2004; Leahy et al., 2005; Leung & Mohan, 2004; Moss, 2008). Informal formative assessment uses everyday learning activities as potential assessments that provide evidence of students' learning in different modes (Eisenkraft, 2004): (a) oral evidence (e.g., students' questions and responses, listening to what they say in small groups, having conversations with students), (b) written evidence (e.g., notes in science notebooks), (c) graphic evidence (e.g., drawing, graphs, drawing concept maps), (d) practical evidence (e.g., observation of students conducting an experiment and measure the mass of an object), and (e) non-verbal evidence (e.g., body language, body orientation). The focus of the assessment can vary from conceptual understanding (e.g., does the student understand relative density?) to practical (e.g., can the student measure the mass of an object using a triple-beam balance?) to social (e.g., are all students participating in the conversation?). Information can involve one or more sources of evidence (e.g., students' questions, students' oral responses, students' written responses in a handout, and student-to-student conversations) and can be accumulated over time before a clear picture of the students' level of understanding can be constructed. For example, the evidence can be accumulated within a class period (e.g., when the teacher uses all the questions and responses asked and given by a particular student across the same class

E-mail address: [Maria.Ruiz-Primo@ucdenver.edu](mailto:Maria.Ruiz-Primo@ucdenver.edu).

period), or over a series of class periods (e.g., different occasions and sources are used to make a decision).

Informal formative assessment can take place within the context of any student–teacher and student–student interaction. It is generally characterized by the conscious discovery of novel information about student understanding, as observed through the context of such interactions at a given point in time. To complete an informal formative assessment cycle, the newly acquired information must be used to shape the immediate course of events within the given learning context. For example, a comment from a student can trigger an assessment event by making the teacher aware of that student’s misunderstanding of a concept. Acting in response to this evidence is usually quick, spontaneous, and flexible, because it can take on different forms (e.g., responding with a question, asking other students to express their opinions, or conducting a demonstration). Relative to formal formative assessments, the timeframe for interpreting and using that evidence is rather limited (Ruiz-Primo & Furtak, 2006, 2007). Therefore, often times, informal formative assessment activities go unrecorded.

Informal formative assessment, then, does not focus on conventional assessment means to collect and interpret information about students, but it is consistent with the purposes of sound educational assessment: it supports instructional decisions based on inferences made about students. Teacher’s decisions are frequently made “on the fly” and allow instruction to move in different directions depending on the ways in which students respond. When informal formative assessment takes place, such “moments of contingency” (William & Leahy, 2007) arise continuously because teachers are constantly seeking to make sense of students’ responses, actions, comments, and behaviors. Conceptualizing formative assessment as collecting information continuously through unconventional means, in addition to conventionally formal (and often less frequent) means, broadens the meaning of the term *assessment*, since this conceptualization does not necessarily rely on associating the term with any particular assessment instrument or task (Moss, 2008). This conceptualization thereby effectively renders assessment as “a normal, ubiquitous part of all social interaction” (Jordan & Putz, 2004, p. 346).

The learning goals that are the focus of the informal formative assessment practices tend to be discrete and immediate (e.g., what students need to get from a particular activity or from that day’s class). However, these discrete, immediate goals are *contextualized* in the larger picture of long-term outcomes; they are part of a *larger learning trajectory* in the context of unit and yearly goals. The cumulative accomplishment of these discrete learning goals is essential if students are to achieve unit or module learning goals. If a teacher makes such discrete learning goals explicit, they can provide a useful context for both interpreting information collected about student understanding and informing teaching decisions within a given lesson. Research on teachers’ pedagogical thinking has clearly indicated that teachers engage in five different levels or “scales” of planning: yearly, term, unit, weekly, and daily (Clark & Yinger, 1987; Shavelson & Stern, 1981), and that planning decisions made early in the academic year have a profound influence on the everyday lesson planning (Shavelson & Stern, 1981). Therefore, the term *informal* does not imply a focus on the naturally unpredictable events that arise in any classroom, but rather on the small-scale, frequent opportunities teachers have for collecting information about their students’ progress towards the learning goals they have in mind. Lampert (2001) describes the different aspects of her students’ understanding that she would like to know about in the scope of preparing multiplication activities (see Moss, 2008 for a careful analysis of this description):

It would be helpful to know whether any of my students could already multiply large numbers using the conventional procedure. It would help to know if anyone understood that multiplication is about groups of groups of groups, no matter how big the number (p. 108).

The informal formative assessment strategies and how a teacher chooses to synthesize and analyze the information gathered through them are critical to ensuring students’ progress towards learning goals. These strategies should be<sup>1</sup>:

- a. *coherent* between each other and with the characteristics of the instruction and the curriculum being implemented; they should also reflect a good understanding of how students develop more sophisticated ways of thinking as the implementation of the curriculum progresses;
- b. *comprehensive*: they should consist of a sufficiently varied collection of pieces of evidence supporting instructional decisions;
- c. *continuous*: they should produce evidence about students’ progress through every opportunity available minute by minute, daily, by lesson, weekly, quarterly, or yearly (Heritage, 2010; Leahy et al., 2005).

#### *Importance of informal formative assessment in the teaching and learning process*

Jordan and Putz (2004) argued that informal assessment constitutes one of the fundamental mechanisms by which learning occurs.<sup>2</sup> Informal assessment can be viewed as a socially situated activity (Jordan & Putz, 2004; Moss, 2008) which extends beyond the classroom and can be observed in a wide variety of contexts, including at home with family and in the work place (Jordan & Putz, 2004). Because of the social nature of classroom activities, a great deal of the information collected through informal formative assessment is through classroom conversations.

Informal formative assessment, then, provides conditions for evaluating and re-thinking the learning opportunities intended to enable students to achieve their learning goals. Evidence collected through informal venues can help teachers to identify patterns in which students talk and their disposition to reflect on what they are doing (see Lampert, 2001).

For informal formative assessment to have an impact on the teaching process and, ultimately, on student learning, it is important for the teacher to adopt a mindset characterized by a desire to (Moss, 2008): (a) continuously learn about students’ learning (e.g., look at variations in the responses of the students, the ways in which they solve problems, or the kinds of errors they make); (b) routinely use this evidence to decide what to do next (e.g., how to respond to students’ misunderstanding, what demonstration can help students to test their hypotheses, or who to call on next); and (c) consider how the instructional practices are shaping students’ learning (e.g., how the tasks selected support students’ understanding and guide them about what is important to learn).

<sup>1</sup> The terms *coherent*, *comprehensive*, and *continuous* were proposed by Pellegrino et al. (2001) in the context of *assessment systems*; here these terms are adopted to reflect the importance that these characteristics also have at the classroom level.

<sup>2</sup> Jordan and Putz (2004) use two terms to talk about informal: *inherent* assessments as implicit, informal and nonverbal (actors choose what to do next), and *discursive* assessments as explicit when members of a social group talk about what they are doing in an evaluative way. They use the term *documentary* assessments as those formal and standardized measurements.

## Instructional dialogues and assessment conversations: the two sides of the same coin

This section focuses on an activity that constitutes a large part of classroom interactions (Bellack, Kliebard, Hyman, & Smith, 1966; Edwards & Westgate, 1994)—*instructional dialogues*, verbal encounters between either the teacher and the students or among the students. Instructional dialogues have an important impact on the process of learning if they are applied to two different perspectives of the same classroom activity, verifying and clarifying student understanding (Bellack et al., 1966; Edwards & Westgate, 1994; Jordan & Putz, 2004). Simply put, an effective dialogue is an interchange of ideas that is “circular in form, cooperative in manner, and constructive in intent” (Martin, 1985, as cited in Roehler & Cantlon, 1997, p. 10).

In the context of informal formative assessment, instructional dialogues can be thought of as *assessment conversations*. Assessment conversations were proposed as a pedagogical strategy by Duschl and Gitomer (1997) as dialogues linked to embedded assessment activities. Later, assessment conversations were identified as an informal formative assessment practice, not necessarily linked exclusively to embedded assessments (Ruiz-Primo & Furtak, 2004, 2006, 2007). Thus, assessment conversations are conceived of as dialogues that embed assessment into an activity already occurring in the classroom. These conversations help teachers to continually acquire information about the level of their students' understanding on the topic at hand in order to shape the instructional activities that can potentially help them to achieve learning goals. The main purpose of assessment conversations is to make students' thinking evident, or to voice their understanding so that teachers can recognize and act on it to promote learning. Assessment conversations make evident *what* and *how* students are thinking, enabling teachers to recognize their students' conceptions, mental models, strategies, language use, and/or communication skills. Teachers can then use this information accordingly to guide the next activities.

### Conceptualizing instructional conversations as assessment conversations

Usually, the study of instructional dialogues in the classroom follows, either partially or completely, the four types of moves proposed by Bellack et al. (1966): *structuring* – moves that set the context for subsequent moves (e.g., attention on a topic or a problem), *soliciting* – moves that elicit verbal or physical responses (i.e., all teacher questions are considered solicitations), *responding* – moves that are the reciprocal to soliciting (e.g., responses from students), and *reacting* – moves that serve to modify (clarify, synthesize, or expand) or evaluate a previous move. These moves have been studied from two very different theoretical perspectives (see Carlsen, 1991). One examines the moves separately (e.g., studying the quality of questions posed by teachers – a soliciting move); the other examines the four moves to identify interaction patterns in the social context of the classroom.

Among the most well-known interaction patterns identified in the classroom are Initiation, Response, Follow-up (IRF) (Sinclair & Coulthard, 1975) and Initiation, Response, Evaluation (IRE) (Mehan, 1979). Most notably, these two patterns have been viewed as limiting the dialogue between teacher and students, due to its similarity with the typical authoritative tone of interaction set by the teacher's pursuit of the “correct response”. Furthermore, both patterns reflect the predominance of teacher talk in the classroom. These well-known interaction patterns may be considered appropriate in the context of assessment conversations. As Wells (1999) pointed out, the triadic dialogues (IRF and IRE; Lemke, 1990) are not intrinsically good or bad; it is the

purpose by which the third move (Follow-up or Evaluation) is made that determines its usefulness for developing conceptual understanding. This remark leaves the door open for interactions of this kind to be productive dialogic interactions (Scott, Mortimer, & Aguiar, 2006), and therefore, to potentially play the role of assessment conversations.

Since the Bellack et al. (1966) study, it was concluded that instructional moves come in teaching cycles which may have the appearance of the four moves discussed before. For example, a cycle can start with an initiation move if the structure move was previously set. There is evidence that assessment conversations sometimes have this circular form and involve *chains of cycles* which establish a sustained interaction between the teacher and the students (Mortimer & Scott, 2000, 2003).

Three issues are important to highlight with respect to assessment conversation cycles. First, an assessment conversation does not have to be initiated by the teacher; a student's comment or question can trigger a cycle (although we know much less about students' contributions than about teachers' contributions to the interaction cycles (however, see Chin, 2002, 2004; Chin & Osborne, 2008; Louca, Tzalli, & Zacharia, 2008).

Second, in addition to teachers, students and peers can act upon or use the information collected depending on the interaction patterns present in the classroom. Teachers as facilitators and orchestrators of the dialogic interactions can allow other students to carry out a strategy that can close the assessment conversation cycle (e.g., providing feedback) or guide a group of students during a discussion to make a joint decision.

Third, assessment conversations can have multiple iterations before the cycle is fully completed or closed. This means that incomplete assessment conversation cycles contribute to facilitating productive thinking. Ruiz-Primo and Furtak (2006, 2007) proposed an alternative interaction pattern – the ESRU cycle – in which the Teacher Elicits information, the Student Responds, the Teacher Recognizes, and the Teacher Uses the information collected to move student learning forward. They observed that ESRU cycles with multiple iterations of incomplete interactions might occur before a dialogue was closed (e.g., ESR → ERS → ERS → ERS → ESRU). However, these multiple iterations tended to close a discussion with a better understanding from the student about the topic at hand. Similar patterns (e.g., IRFRF) have been cited in other studies (Mortimer & Scott, 2003). Ruiz-Primo and Furtak (2006, 2007) also observed that students whose teachers implemented greater numbers of dialogic interactions that ended the chain with the *act/use* of information performed better on tests on the topics discussed than students whose teachers did not implement those dialogic interactions. Also, Minstrell et al. (2009) found that teachers who are highly competent in implementing formative assessment have more elaborate (cyclical) conversations, with more turns to the students, than teachers who are less competent. Thus, closure is critical to successfully implementing assessment conversation cycles.

### Conditions for effective assessment conversations

Assessment conversations must be re-thought and re-conceptualized by considering strategies that maximize their efficacy. The following set of conditions needs to be met if they are to be an effective, informal formative assessment strategy:

- *Effective assessment conversations are learning goal-guided.* Assessment conversations should not be seen as verbal encounters that happen only in the classroom; they should be conducted with a *learning goal in mind*. Assessment conversations are instructional conversations intended to draw out information from the class, a group of students, or an individual



student in an attempt to better define the course of instruction that can ultimately lead to accomplishing learning goals. Clear learning goals and clarity about what constitutes evidence of having met these goals are critical in successfully linking instructional and assessment practices.

- *Effective assessment conversations are dialogic and interactive in nature.* In a dialogic interaction, the teacher encourages, recognizes, and attempts to take into account a range of students' ideas to have a clear picture of students' understanding (Scott et al., 2006). In dialogic interactions, students' responses are seen as contributions that can be discussed by their teachers and peers. Assessment conversations should allow students to argue and justify their ideas, engage in considering each other's points of view, foster the use of evidence, and facilitate the discussion on the quality of this evidence to define the appropriateness of their claims (Chin, 2006, 2007; Duschl & Gitomer, 1997; Duschl & Osborne, 2002; Hogan & Pressley, 1997; Scott et al., 2006). In a non-interactive dialogue, there is no turn-taking with students; the teachers address student responses by shaping or reviewing their ideas (Mortimer & Scott, 2003).
- *Effective assessment conversations are applied as instructional scaffolding tools.* Assessment conversations should provide a balance between challenging and supporting students in order to help them achieve discrete learning goals on a daily basis. Inviting student participation, verifying and clarifying students' understanding, and offering examples, explanatory reasoning, and elaborations are examples of the types of scaffolding that occur in successful classrooms (Roehler & Cantlon, 1997). Assessment conversations allow for exchanges between more experienced, expert members (the teacher and highly proficient students) and less experienced, novice members (less proficient students) that can reduce the cognitive distance between what students currently know and can do, what they are capable of knowing and doing with assistance from experts, and what they can learn to know and do on their own (Hogan & Pressley, 1997). Assessment conversations, if successful, should both support students in thinking for themselves – which promotes self-regulation (Hogan, 1997) – and support their ability to reason and reflect metacognitively on their own learning (Bransford, Brown, & Cocking, 1999; Kuhn, 1999).
- *Effective assessment conversations are supportive tools of social participation and social cognition.* Assessment conversations should occur in an environment that encourages students to ask questions and to justify and evaluate each others' reasoning and that supports them in constructing knowledge through reasoning. This type of environment promotes higher-order cognitive skills – skills that go beyond declarative knowledge to promote schematic knowledge (see Li, 2001; Li, Ruiz-Primo, & Shavelson, 2006) and metacognitive practice (*knowing how to learn*; see Bransford et al., 1999; Kuhn, 1999).
- *Effective assessment conversations are enculturation tools.* Assessment conversations should serve to immerse students into the language, culture, and artifacts of the academic discipline. For example, in the context of science education, this environment should allow students to engage in dialogues that resemble those of the scientific community (Duschl, 2003; Duschl & Osborne, 2002) by promoting more sophisticated ways to reason and communicate (e.g., weighing evidence, generating hypotheses, evaluating theories, and communicating ideas).

Assessment conversations are effective when teachers have sufficient content knowledge and pedagogical content knowledge to interpret and act spontaneously and contingently on students' contributions. Only if they are equipped with deep content knowledge can they recognize evidence of students' level of understanding (e.g., students' misunderstandings) and address,

through different strategies, the issues found in the information collected about students' learning. Minstrell et al. (2009) found that skilled teachers have deep understanding of the content knowledge. This knowledge allows them to: (a) be more explicit, responsive, and nuanced in their interpretation of student responses and in their feedback to students; (b) be able to anticipate the most critical weaknesses and strengths in student performance; (c) anticipate more than one problematic response; (d) interpret student responses based on scientific principles rather than on superficial responses; and (e) better determine the instructional actions they need to take in order to approach the central issues associated with student needs and in order to provide descriptions that are detailed enough so as to go beyond superficial characteristics.

#### Assessment conversations strategies

The description of the assessment conversation strategies follows the assessment activity cycle proposed by Bell and Cowie (2001): collecting information, interpreting information, and acting-on/using the information collected. A fourth assessment activity not mentioned by these authors was added as the first activity of the cycle (Ruiz-Primo et al., 2007): clarifying the learning expectations (goals) as a prerequisite for collecting, interpreting, and acting on the information with the intention of improving student learning. Furthermore, clarifying the learning expectations should involve determining what will be considered as evidence of achieving those learning expectations or goals (Sadler, 1989).

*Clarifying learning goals.* Assessment conversations take place only when there is a clear learning goal guiding the interactive dialogues and there is clarity of the kind of evidence needed to know that the learning goal has been met. If such reference framework does not exist, the conversations become ineffective for assessment purposes. As mentioned earlier, learning goals can be considered at different levels: by course, by unit or module, lesson, session, or activity. Teachers can explain the learning goal or the learning target as well as the success criteria to the students; they can involve students in discussing the learning target and developing success criteria. Harlen (2007) suggests that communicating the learning goals to students should include two aspects: (a) explaining what students will be learning *about* (e.g., relative density), and (b) explaining *how* they will be learning (e.g., testing with objects of different weight and size in different liquids).

Clarifying the learning goals and revisiting them multiple times help students to stay focused and motivated and to value what they are doing. Minstrell et al. (2009) found that teachers who are skillful in implementing formative assessments introduce, remind, and discuss learning goals three times more frequently than less skillful teachers. Obviously, learning goals cannot be shared and discussed with students during every assessment conversation. However, a learning goal or a target that is in the teacher's mind has the potential to be included as part of an assessment conversation through the following actions:

- *Reminding the students about the learning goal or learning target during the conversation* (e.g., "Since we want to learn about why things sink or float, what questions can we ask about these objects?" or "Keeping in mind how we conducted the investigation of weight in terms of floating and sinking, we will attempt to do the same investigation of volume. Do you remember what we hope to learn by doing this investigation?").
- *Reminding students about the purpose of an activity* (e.g., "Why do you think it is important to weigh these objects and submerge them in water?" "What do you think we will learn by doing this activity?").

- *Connecting the discussion (conversation) to the learning goal* (e.g., “We are discussing our results about the weight of these objects because we are still trying to figure out why things sink or float.”).

*Collecting/eliciting information: questioning.* The strategies used to collect information that makes students’ thinking evident to the teacher are critical in determining the nature of the dialogue. *Questioning* is a strategy that can be easily used to start and continue an assessment conversation. The kinds of questions used, as well as the manner in which they are asked, can influence the type of cognitive processes in which students engage (Chin, 2007; DeWitt & Hohenstein, 2010; Harlen, 2007; Minstrell & van Zee, 2003; van Zee & Minstrell, 1997). Questions can be of different types and serve different purposes. A distinction can be made between questions posed by the teacher in a deliberate act looking for evidence of students’ understanding – “purposive evidence” (William & Black, 1996) – and questions posed by the students. Such student-generated questions may provide teachers with “unexpected opportunities” to gather information about students’ understanding – “incidental evidence” (William & Black, 1996). Questions posed by teachers have been categorized in different ways. For example, in the context of science education, questions have been classified on a continuum ranging from those that cannot stimulate productive thinking (e.g., questions to adjust teacher’s agenda; Chin, 2007) to those that can stimulate productive thinking (e.g., Socratic questions, verbal jigsaw, semantic tapestry, and framing questions) to those tapping into different scientific dimensions (e.g., epistemic, methodological, conceptual, and social; Ruiz-Primo & Furtak, 2006, 2007).<sup>3</sup>

In assessment conversations, it is important to consider the adaptability required in questioning (Chin, 2006, 2007; Ruiz-Primo & Furtak, 2006, 2007). For example, sometimes questions are used to *elicit information* (e.g., “Why do you think this goes with this?”) or to *invite student participation* (e.g., “Who can expand on what Jeanne said?”); other questions focus on *building upon a student’s response* (e.g., “Why do you think so?” “How do we know this is a mixture?”); still others are a response to a question posed by a student to throw the responsibility of thinking back to the student—a reflective toss such as, “I do not know, what do you think?” (see van Zee & Minstrell, 1997). Teachers adjust questioning to accommodate students’ contributions and to respond to students’ thinking according to the assessment move (e.g., collecting, interpreting, and using). The role of questions in assessment conversations can only be understood when there is an understanding of the context in which they occur (e.g., “What is the assessment move in which the question occurs?” “What is the student response?” (see Carlsen, 1991; DeWitt & Hohenstein, 2010). Teachers’ effective questioning provokes cognitive discrepancies that promote more complex cognitive activities (Duschl & Gitomer, 1997; Hogan & Pressley, 1997). For assessment conversation purposes, teachers’ questions should be:

- *Open-ended.* This type of question invites students to probe their thoughts and to generate ideas based on reasoning and prior knowledge (Chin, 2007; DeWitt & Hohenstein, 2010; Minstrell & van Zee, 2003; van Zee & Minstrell, 1997). Open-ended questions

should be phrased in ways that can better elicit what students think (e.g., “What experiment *do you think* can be done to test this hypothesis?”; Harlen, 2007) instead of focusing on questions that seem to allude to a particular response (e.g., “What experiment can be done to test this hypothesis?”).

- *Tapping into diverse types of knowledge.* Questions can tap into the following four types of knowledge (Li, 2001; Li et al., 2006): declarative knowledge (knowing what), procedural knowledge (knowing how), schematic knowledge (knowing why), and strategic knowledge (knowing why, when, and how to use the knowledge).<sup>4</sup> Questions such as “Why do you think so?”, “How do you know that?”, or “What evidence do you have to support your claim?” more directly elicit what students’ thinking is (Ruiz-Primo & Furtak, 2006, 2007) and further stimulate students’ thinking (Chin, 2007). However, questions tapping into declarative (e.g., “Who can tell me what is density?) or procedural knowledge (e.g., “What did you do to mass the cork?”) should be pursued if the teacher’s focus is on the use of, for example, the appropriate language within a discipline (enculturation). Minstrell et al. (2009) found that teachers highly skilled in implementing formative assessment practices use a variety of question types to probe students’ thinking. They also found that these questions tend to tap into more than one knowledge type, whereas less skillful teachers tend to use questions that tap procedural and, to a lesser degree, declarative knowledge only, and rarely formulate questions that tap into knowledge outside of those realms.

As mentioned, students’ questions can also provide the “incidental” opportunity for a teacher to conduct an assessment conversation and gather evidence about their students’ level of understanding. Students’ questions are considered potential resources for teaching and learning (Chin & Osborne, 2008; Scardamalia & Bereiter, 1992). It has been argued that, in order for student questions to emerge, a teacher must deliberately attempt to stimulate them (Chin, 2004; Chin & Osborne, 2008). One teacher strategy aimed at fostering student questions involves providing students with the opportunity to work in small groups, which provide the context for social engagement and shared tasks. Alternatively, the teacher can directly model the development and use of questions (Chin, 2004; Chin & Osborne, 2008; Mercer, Dawes, Wegerif, & Sams, 2004). However, student questions can also be posed in a lecture or as part of an assessment conversation when more than one student poses questions as a result of a teacher’s or student’s contribution (adjacent assessment conversations). There are interventions that have been developed to improve student questioning (e.g., Chin, 2004; Mercer et al., 2004). In the context of informal formative assessment, it is most instructive to think about the nature of the questions students are asking in terms of how they may contribute differently to the quality of the assessment conversation. Some student questions may be more superficial than others. The following aspects can be used to distinguish between the nature of student’ questions (see Chin, 2004; Chin & Osborne, 2008):

- *Type of information asked in the question.* The content of the student question can focus on: (a) procedural information (e.g., paying attention to what students are doing; “How long should my response be; three, four, ten lines?”); (b) curiosity (e.g., “Is it known whether all types of wood float?”); (c) factual information (e.g., “What is the density of alcohol?”); (d) clarifications (e.g., question looking for more detail; “Do we always weigh the object to find out the density?”); (e) confirmation or consolidation (e.g., “So, objects with mass identical to water, will they always float?”); (f) explanations (e.g., “Why doesn’t this wood cube sink?”; “Do all wood objects float like the boats in the river?”);

<sup>3</sup> These dimensions of science education were originally proposed by Driver et al. (1996), elaborated by Duschl (2000, 2003), and informally revised by Ruiz-Primo and Furtak (2006) and formally by Furtak (2006).

<sup>4</sup> Declarative knowledge focuses on factual and conceptual knowledge. Procedural knowledge focuses on algorithms or step-by-step or condition-action sequences. Schematic knowledge refers to knowledge used to reason about, predict, and explain things in nature; it focuses on student’s mental models. Strategic knowledge (knowing when, where, and how knowledge applies) focuses on deep understanding of knowledge that allows students to use in new situations and contexts.

(d) transformations—look for extensions of what is known, identify omissions, challenge some reasoning; examine claims (e.g., “We have been studying objects that sink or float. What about objects that do not completely sink and do not completely float?”).

- *Depth of understanding reflected.* Some questions reflect a deeper understanding than others. Those questions that reflect curiosity or knowledge-base knowledge and wonderment can more positively affect the learning process (Scardamalia & Bereiter, 1992).
- *Type of engagement during questioning.* The type of question asked may reflect the type of engagement in which the students are involved: Procedural engagement is considered as the type of engagement that reflects that the student is paying attention and does the work but is not substantially committed to the content being discussed. Questions focusing on the content being studied reflect a substantive engagement (Nystrand & Gamoran, 1991).

*Interpreting information collected.* In assessment conversations, making sense of students’ responses (the information collected) happens *on the fly* and often times the interpreter is the teacher, although students can be interpreters too. The detailed analysis and interpretation of the students’ responses that take place in a formal formative assessment context are less likely to happen in an informal context. Sometimes, the interpretation made by the teacher is not evident to an external observer. In assessment conversations, teachers need to be careful in judging openly students’ responses or explanations of their thinking. For students to feel confident and safe in their classrooms, their contributions (e.g., their responses, ideas, and explanations) should be carefully considered to avoid embarrassment (Chin, 2004; Hodgson, 2010). However, this does not mean that the teacher is not making sense of the students’ responses based on the learning target the teacher has in mind. Different strategies have been suggested for teachers to value or to take into account students’ responses without judging them directly:

- *Rephrasing, clarifying, elaborating, summarizing, and repeating* help to verify students’ responses. Verifying student contributions acknowledges student effort. It signals knowledge as a critical component in instructional dialogues and provides teachers with another opportunity to prompt students’ active reflections that may in turn help them to refine their thinking. Minstrell et al. (2009) found that teachers highly skillful in implementing formative assessment paraphrased students’ contributions to verify their thinking or summarize their contributions, proportionally, five more times than less skillful teachers.
- *Relating a student’s response to another student’s response.* The teacher incorporates the student’s contribution into the conversations by referring to another student’s contribution (e.g., “What you are saying is similar to what Becca mentioned before. . .”). This strategy has been named *uptake* (Nystrand & Gamoran, 1991)—a teacher “takes up” the students’ previous response or comment and builds it into the next question to create continuity in the discourse.
- *Promoting students’ questions and comments about a student’s contribution.* The teacher encourages students to ask questions or to make comments about what was just said (e.g., “John, what do you think about what Karla just said?” “Does anybody want to add something to what Karla just said?”).
- *Displaying students’ responses.* This strategy not only helps teacher to value students’ responses but it also helps students to learn from each other. For example, students’ responses can be displayed on chart paper or on white boards for a plenary discussion (e.g., how the responses can be grouped according

to certain criteria that can be proposed by the teacher or by the students or by both (Chin, 2004; Hodgson, 2010; Ruiz-Primo & Furtak, 2006, 2007). Allowing students to compare and analyze their responses constitutes an example of the students being the interpreters and the teacher facilitating such interpretation.

- *Responding with a reflective toss.* Reflective toss is the response provided by teachers to “throwing” the responsibility for thinking to the students (van Zee & Minstrell, 1997). A reflective toss requires the teacher “catching,” making sense of the student’s prior question or statement, and responding with another question that the student will respond. This strategy provides the responsibility to think not only to the students participating in the assessment conversation but also to other students in the class (van Zee & Minstrell, 1997).

*Acting on the information collected.* In this section, different strategies are proposed as ways to act on the information collected. This assessment activity is rarely enacted in the classroom (Ruiz-Primo & Furtak, 2006, 2007).

- *Offering explanations.* This strategy has been seen as a critical scaffolding strategy. Explanations are statements about what is being learned that are adjusted to fit the level of the students’ understanding. Explanations which better serve scaffolding purposes provide information about *why* something is important, *when* it is used, and *how* it is used (Gaskins et al., 1997; Lepper, Drake, & O’Donnell-Johnson, 1997; Roehler & Cantlon, 1997). This type of explanation contributes to the development of metastrategic processes (i.e., knowing which strategies to deploy; see Kuhn, 1999).
- *Comparing and contrasting students’ ideas/contributions.* This strategy directly uses the collected evidence to clarify students’ thinking. Comparing and contrasting students’ ideas promote peer assessment where students listen to the ideas of others, consider supporting evidence of the statements given, and help to progress to higher levels of understanding (Sadler, 1989). Comparing and contrasting students’ ideas can lead to the construction and evaluation of explanations based on the evidence available. In the context of science, this strategy is thought of as a way to engage students in argumentation – “critically exploring the coordination of evidence and theory that support or refute an explanatory conclusion, model or prediction” (Duschl & Osborne, 2002, p. 44). Argumentation is seen as a central strategy which stimulates the process of reflection that lead to students’ conceptual understanding (Duschl, 2003; Duschl & Osborne, 2002; Lawson, 2003).
- *Providing feedback.* The recent attention to feedback in the context of formative assessment leaves few things to add (e.g., Black, Harrison, Lee, Marshall, & Wiliam, 2003; Brookhart, 2004, 2007, 2008; Cameron & Pierce, 1994; Hattie & Timperley, 2007; Kluger & DeNisi, 1996; Shute, 2008; Torrance & Pryor, 1998; Yeany & Miller, 1983). We know that effective feedback should lead students to be able to judge the quality of what they are producing and to be able to monitor themselves during the act of production (Sadler, 1989, 1998). We know that effective feedback should lead to answer the questions, where am I going? How am I doing? And where to go next? (Hattie & Timperley, 2007; Ramaprasad, 1983). We know that teachers who are skillful in implementing formative assessment provide feedback that is more frequent and more appropriate to student needs, proportionally, about three times more than less skillful teachers (Minstrell et al., 2009). When quality feedback is provided by teachers and/or peers, students are empowered to take the appropriate action which leads to improved self-regulation (Andrade, 2010). To effec-



tively assess the quality of students' work or performance, teachers and peers must possess a concept of quality appropriate to the task that allows them to recognize and describe quality performance and indicates how poor performance can be improved (Sadler, 1989). Two more issues in the context of assessment conversations can be added to the discussion of feedback. First, according to the nature and quality of dialogic interactions, there are occasions in which *assessment conversations can be considered by themselves as a feedback strategy* in which feedback is indirectly provided to a non-participating student. Evidence from observations suggest that a student can reach new levels of understanding by witnessing teachers interact with other students when some sort of feedback or evaluation transpires (Bandura, 1977; Bangert-Drowns, Kulik, Kulik, & Morgan, 1991). Assessment conversations can help connect students' own ideas to others' ideas. This notion is aligned with the view of scaffolding as a "weaving metaphor" that expands beyond dyadic interactions into whole classrooms (Hogan & Pressley, 1997). Second, in assessment conversations, it is important to focus not only on the "cognitive" feedback but also on the "affective" feedback. Different types of feedback affect students' beliefs about their performance capabilities and the value of effort in producing good performance, which in turn affects student motivation (Schunk, Pintrich, & Meece, 2007; Stipek, 1984). For example, attributional feedback that addresses students' self-efficacy or effort is as important as the cognitive feedback (e.g., Brookhart, 2008; Butler & Neuman, 1995; Hogan & Pressley, 1997; Schunk, 1982; Schunk & Swartz, 1993; Schunk et al., 2007; Shute, 2008). Not enough emphasis has been put on this type of feedback in formative assessment (see Shute, 2008). We know that expert tutors who focus on cognitive and motivational/affective feedback increase not only students' learning, but also students' interest and enthusiasm in the study materials (see Lepper et al., 1997).

- *Modeling*. This strategy is a scaffolding activity that may involve different possibilities: demonstrating to students the thought process (thinking-aloud) underlying the steps in a task, showing by both acting and talking through the steps of the task, or performing the task (Roehler & Cantlon, 1997). Modeling with talking aloud makes thinking visible to students and encourages them to do the same in approaching a task. As with explanations,

teachers can explain *what* strategy they are using, *why* the strategy is the appropriate one, *how* to implement it, and *when* it can be used.

- *Debugging*. This is a strategy used by expert tutors when students have made errors (or are permitted to make particular errors) and are guided, through hints and questioning, to identify and correct themselves (e.g., "So how did you get that number?" "Which number do you really want to use?" "How did you arrive at this answer?" Lepper et al., 1997). When debugging fails, modeling can be used. Usually, expert tutors start by posing very general and indirect questions to the students to avoid direct confrontation of the error.

### Empirical evidence on the effect of assessment conversations on student learning

Since instruction normally involves "verbal encounters" orchestrated by teachers (Edwards & Westgate, 1994), studying assessment conversations necessarily involves investigating classroom talk. Many studies focus on classroom interactions (almost all of the citations used in this paper focus on classroom interactions). However, fewer studies exist that provide evidence about the link between the quality of those interactions and student learning. Many studies on classroom interactions use small numbers of cases (teachers), or even a single case in fine-grain detail (e.g., Chin, 2006; Duschl & Gitomer, 1997; Hogan & Pressley, 1997; Leung & Mohan, 2004; Louca et al., 2008; Mehan, 1979; Roehler & Cantlon, 1997; Roth, 1996; Ruiz-Primo & Furtak, 2006, 2007; Scott et al., 2006; van Zee & Minstrell, 1997). While this raises concerns about the generalizability of the findings, it can be argued that these studies not only shed some light on the process of interaction, but also that collectively they may provide some evidence about the consistency across cases in different studies.

Studies looking at classroom interactions with large samples of observations are scant due to the amount of work needed to collect data. Therefore, very few studies are presented in this section that provide some evidence about the role of instructional dialogues, studied as interaction patterns, and assessment conversations. Table 1 provides a general picture of some of these studies. The intention is not to present a meta-analysis or provide a literature review of studies conducted in this area, but to point at characteristics that future studies should consider.

**Table 1**  
General characteristics of studies linking student learning and instructional dialogues.<sup>a</sup>

Study	$n_T$	$n_S$	$n_O$	Grade	Coding reliability	Design	Learning outcomes	Technical qualities of outcome measures	
								Validity	Reliability
Bellack et al. (1966)	15	345	4	High school	High – Adjudication <sup>b</sup>	Pre-Post, no control	– Knowledge: 3 essays – Attitudes	Content <sup>c</sup>	IRR = .84
Nystrand and Gamoran (1991)	58	1041	4	8th Grade	High – Adjudication <sup>b</sup>	Post, no control	– Literature achievement test	Content <sup>c</sup>	IRR = .90
Applebee et al. (2003)	64	1111	4	Middle and High-school	No information provided	Students measured on four occasions: at the beginning and on 3 more occasions	– Task 1: Open-ended on why a character is admired. – Task 2: Essay – Task 3: Essay	Content <sup>c</sup>	High – adjudication <sup>b</sup>
Ruiz-Primo and Furtak (2006)	4	90	12 <sup>a</sup>	7th Grade	.86	Pretest and 3 embedded assessments	1. Multiple-choice (Pre) 2. Graphing 3. Predict–observe–explain 4. Predict–observe	Content <sup>c</sup>	1. IC = .86 2. IRR = .97 3. IRR = .98 4. IRR = .94
Ruiz-Primo and Furtak (2007)	3	76	10 <sup>a</sup>	7th Grade	.74	Pre–post and embedded assessments	Multiple-choice (pre–postest)	Content <sup>c</sup>	IC = .86

<sup>a</sup>  $n_T$  = number of teachers;  $n_S$  = number of students;  $n_O$  = number of observations; IRR = inter-rater reliability; IC = internal consistency.

<sup>b</sup> The source mentions that an agreement check was used as well as adjudication when disagreement was found. However, no specific percentage of agreement or inter-rater reliability coefficient is provided.

<sup>c</sup> The source mentions that the measure was developed after a careful analysis of the material and/or curriculum to be studied by the students.

At first glance, it is clear that some of these studies used a large number of teachers/classrooms. The size of the student samples in the first three studies are large enough to allow for sophisticated statistical analyses. In these studies, four classrooms observations were conducted. A difference with the last two studies, with much smaller sample of teacher and students, is that at least 10 classroom sessions per teacher were analyzed. In all studies, time was considered a critical issue; either by sampling over time or continuously tracking teachers' instructional dialogues. In all studies, the moves proposed by Bellack et al. (1966) – whether named soliciting, initiating, questioning, or eliciting – guided the analyses in one way or another. In all studies, the nature of the questions involved in the interactions played a critical role in the analysis, whether the questions were posed by the teacher or by the students. Another critical aspect is how the teacher followed up to the students' response – how they acted based on the information collected. Finally, in all of the studies the focus was the interaction patterns, rather than individual triads (IRE or IRF).

All the studies showed a positive effect of desirable classroom interaction patterns with students' learning outcomes. For example, Bellack et al. (1966) observed the highest percentage of structuring moves (such as those focusing attention on the topic) but the lowest percentage of soliciting moves focusing on fact-based information among the three highest performing classes. Using regression analysis, Nystrand and Gamoran (1991) found that students in classes with more uptake, more authentic questions (questions with no pre-specified answers), and more discussion time, obtained high scores on the achievement test. Disengagement had a strong negative impact on achievement. Applebee, Langer, Nystrand, and Gamoran (2003) found similar results. Using hierarchical linear modeling, they found that classroom interactions that reflected a discussion approach were positively related to the students' performance. The positive impact was observed in both high- and low-achieving students. Finally, Ruiz-Primo and Furtak (2006, 2007) provided evidence that linked the use of more assessment conversations with higher student learning. Their results indicated that the teachers whose enactment of informal formative assessment was more consistent with the ESRU model had students with higher performance on embedded assessments. This trend was also reflected in the post-test scores.

### **Moving forward our knowledge on informal formative assessment**

Studying informal formative assessment practices is a complex process. The methods are more costly and time consuming than the actions taken when, say, a test or a questionnaire is administered. Data collection and analysis is costly because videotaping is the only data collection method that can be used if exchanges are to be studied in detail. Classroom observations may not provide the necessary focus to collect this information in real time. Transcribing and developing and implementing coding schemes are also tedious and costly. Therefore, studies with large sample sizes are rarely conducted. How, then, can we move this research area forward? Studies such as those conducted by Nystrand and Gamoran (1991) and Applebee et al. (2003) give us hope that we can increase our knowledge about informal formative assessment practices. We need to learn more about the logistics of conducting cost-effective large-scale studies that focus on small units of analyses as exchanges between teachers and students. We need to explore other forms of analyses that can guide us to learn more about contextual factors that impede or facilitate the type of classroom interactions that we know are more productive in improving student learning. Although continuing research on the details of the classroom interactions is beneficial, we need to start

providing more empirical evidence about what exactly happens to student learning when each particular type of interaction transpires. This information has great potential to speak directly to pedagogical improvements at the classroom level. By now, we know enough about how IRF and IRE dominate interaction patterns in a classroom. The real issue is what we plan to do with that knowledge.

Focusing on other forms of analyses may provide better clues about the conditions that constrain or facilitate the investigation of dialogic interactions, whether or not we want to call them assessment conversations. For example, we should start focusing on *sequential analysis* to conduct systematic observations that allow us to move from merely describing to modeling in order to better explain what happens in the classroom (Backerman & Gottman, 1997; Gottman & Roy, 1990). This will allow research to move from merely tallying frequencies and percentages to calculating transitional probabilities across time, so that event-based probabilities are better understood. Researchers should further explore other forms of analysis recently introduced in educational contexts. For example, in a second study, Nystrand, Wu, Gamoran, Zeiser, and Long (2001) employed *event history techniques* to uncover the conditions that lead to the emergence of dialogic interactions, discussions, and student questions in 200 classrooms. Event history analysis is a recently developed quantitative technique for investigating the causes and results of events. Unfortunately, this study does not link information of dialogic interactions with student learning. According to these authors, this technique has been used for some years in other disciplines but with different names (e.g., survival analysis in biostatistics, failure time analysis in engineering, or duration analysis in economics). The education field has much to gain by using interdisciplinary research teams (as in the study by Nystrand and Gamoran) that allow for more sophisticated analyses and permit the development of models that can provide better explanations of the phenomena of classroom interactions.

Finally, it seems that there are other areas in need of study. We know about certain aspects of questioning (Carlsen, 1991; Cazden, 2001; Chin, 2007; Roth, 1996; Scott et al., 2006) and feedback (Shute, 2008), but we do not know much about how questioning strategies change over time, or how questions are adapted according to students' progress. Also, we do not know much about certain types of teachers' responses to students' errors. What types of dialogic interactions have the largest effect on student learning? What kinds of classroom interactions are related to what kinds of learning outcomes? Is there an interaction between the characteristics of the class and the type of dialogic interaction? Is the effectiveness of dialogic interactions dependant on the type of context? Responses to these questions require larger sample sizes and, probably, more sophisticated data analyses. Although some research has investigated the ways in which the characteristics of learning environments (e.g., tasks, topics) influence student motivation and engagement (e.g., Blumenfeld, Kempler, & Krakcik, 2006), motivational exchanges in the classroom are yet to be studied in detail. As mentioned, affective feedback is as important as cognitive feedback because it affects students' beliefs about their performance capabilities; yet, research has not examined these types of exchanges sufficiently.

### **Conclusions**

Four ideas guided the content of the paper: *how* informal formative assessment can be conceptualized, *how* instructional dialogues can be considered informal formative assessment sources of information, *what* evidence exists regarding its effect on student learning, and *where* we are in our knowledge about informal formative assessment and what we can do to advance it.



Informal formative assessment was conceptualized around the idea that everyday learning activities can be seen as potential assessments that provide evidence of student learning. Almost everything that teachers and students do in classrooms can be described as an opportunity for collecting evidence of their students' understanding. Informal formative assessment is critical for teaching and learning because it makes students' thinking evident. When students' thinking is made evident, it can be examined, questioned, and shaped as an "active object of constructive learning" (Glaser, 1995, cited in Duschl & Osborne, 2002). Knowing where students stand on a day-to-day basis enables teachers to determine where they are in relation to where they should be, so that they can provide the appropriate scaffolding to move their students forward in their learning. The distance between the actual level of understanding and the potential level of understanding that a student can achieve with the help of the teacher aligns with the notion of zone of proximal development within the social constructivist perspective (Vygostsky, 1978). Therefore, the assessment activities proposed in the assessment conversations are a natural way to provide scaffolding to students.

This paper had little to say about *what evidence exists* regarding the effect of informal formative assessment on student learning. Although there are many studies focusing on classroom interactions, few focus on how the quality of these interactions influences students' learning, especially in science. Still, the few that address this link are hopeful—they find a positive impact.

Formative assessment has a long way to go before it is practiced in a masterful manner. There are many items that need to be included in a research agenda on informal formative assessment. Among the most important are those that hold promise to impact teacher practice through teacher education and professional development programs.

## References

- AFL. (2009). Position paper on assessment for learning. *Third international conference on assessment for learning* Retrieved December 21st 2010, from: <http://www.fairtest.org/position-paper-assessment-learning>.
- Andrade, H. L. (2010). Students as the definite source of formative assessment. Academic self-assessment and the self regulation of learning. In G. J. Cizek & H. L. Andrade (Eds.), *Handbook of formative assessment* (pp. 90–105). New York, NY: Routledge. Taylor & Francis Group.
- Applebee, A. N., Langer, J., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal*, 40(3), 685–730.
- Assessment Reform Group (2002). *Assessment for learning: 10 principles*. Research based principles to guide classroom practice. Paper retrieved July 1st 2010, from: <http://www.assessment-reform-group.org.uk/CIE3.pdf>.
- Backerman, R., & Gottman, J. M. (1997). *Observing interaction. An introduction to sequential analysis* (2nd ed.). Cambridge, NY: Cambridge University Press.
- Bandura, A. (1977). *Social learning theory*. Prentice-Hal: Englewood Cliffs, NJ.
- Bangert-Drowns, R. L., Kulik, C.-L., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61(2), 213–238.
- Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht, The Netherlands: Kluwer.
- Bellack, A. A., Kliebard, H. M., Hyman, R. T., & Smith, F. L. (1966). *The language of the classroom*. New York: Teachers College Press.
- Black, P. (2009). From frequent testing to interactive dialogue: Diversity and challenge. Guest's editor introduction. *Journal of Multidisciplinary Evaluation*, 6(12), iii–vi.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.
- Blumenfeld, P. C., Kempner, T. M., & Krakcik, J. S. (2006). Motivation and cognitive engagement in learning environments. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 475–488). New York, NY: Cambridge University Press.
- Bransford, J., Brown, A., & Cocking, R. (1999). *How people learn: Brain, mind, experience and school*. Washington, DC: National Academy Press.
- Brookhart, S. M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record*, 106, 429–458.
- Brookhart, S. M. (2007). Feedback that fits. *Educational Leadership*, 65(4), 54–59.
- Brookhart, S. M. (2008). *How to provide effective feedback to your students*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Butler, R., & Neuman, O. (1995). Effects of task and ego achievement goals on help-seeking behaviors and attitudes. *Journal of Educational Psychology*, 87, 261–271.
- Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research*, 64(3), 363–423.
- Carlsen, W. S. (1991). Questioning in classrooms: A sociolinguistic perspective. *Review of Educational Research*, 61, 157–178.
- Cazden, C. B. (2001). *Classroom discourse. The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.
- Chin, C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education*, 24(5), 521–549.
- Chin, C. (2004). Students' questions: Fostering a culture of inquisitiveness in science classrooms. *School Science Review*, 86(314), 107–112.
- Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to students' responses. *International Journal of Science Education*, 28(11), 1315–1346.
- Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching*, 44(6), 815–843.
- Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1–39.
- Clark, C. M., & Yinger, R. J. (1987). Teacher planning. In J. Calderhead (Ed.), *Exploring teachers' thinking* (pp. 84–103). London, UK: Cassell Educational Limited.
- Daws, N., & Singh, B. (1996). Formative assessment: To what extent is its potential to enhance pupils' science being realized? *School Science Review*, 77(281), 93–100.
- DeWitt, J., & Hohenstein, J. (2010). School trips and classroom lessons: An investigation into teacher–student talk in two settings. *Journal of Research in Science Teaching*, 47(4), 454–473.
- Driver, R., Leach, J., Millar, R., & Scott, P. H. (1996). *Young people's images of science*. Buckingham, UK: Open University Press.
- Duschl, R. (2000). Making the nature of science explicit. In R. Millar, J. Leach, & J. Osborne (Eds.), *Improving science education. The contribution of research* (pp. 185–206). Buckingham, UK: Open University Press.
- Duschl, R. A. (2003). Assessment of inquiry. In J. M. Atkin & J. E. Coffey (Eds.), *Everyday assessment in the science classroom* (pp. 41–59). Arlington, VA: National Science Teachers Association Press.
- Duschl, R. A., & Gitomer, D. H. (1997). Strategies and challenges to changing the focus of assessment and instruction in science classrooms. *Educational Assessment*, 4, 37–73.
- Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation in science education. *Studies in Science Education*, 38, 39–72.
- Edwards, A. D., & Westgate, D. P. G. (1994). *Investigating classroom talk*. London, UK: Falmer Press.
- Eisenkraft, A. (2004). How do we know what they know. *FASS Meeting*.
- Furtak, E. M. (2006). *The dilemma of guidance in scientific inquiry teaching*. Stanford, CA: Stanford University. [unpublished dissertation].
- Gaskins, I. W., Rauch, S., Gensemer, E., Cunicelli, E., O'Hara, C., Six, L., et al. (1997). Scaffolding the development of intelligence among children who are delayed in learning to read. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning. Instructional approaches and issues* (pp. 43–73). Cambridge, MA: Brookline Books.
- Gearhart, M., Nagashima, S., Pfothauer, J., Clark, S., Schwab, C., Vendlinsky, T., et al. (2006). Developing expertise with classroom assessment in K–12 science: Learning to interpret student work. Interim findings from a 2-year study. *Educational Assessment*, 11(3–4), 237–263.
- Gottman, J. M., & Roy, A. K. (1990). *Sequential analysis. A guide for behavioral researchers*. Cambridge, NY: Cambridge University Press.
- Harlen, W. (2007). Formative classroom assessment in science and mathematics. In J. H. McMillan (Ed.), *Formative classroom assessment* (pp. 116–135). New York, NY: Teacher College Press.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–113.
- Heritage, M. (2010). *Formative assessment. Making it happen in the classroom*. Corwin: Thousand Oaks, CA.
- Hodgson, C. (2010). *Assessment for learning in primary science: Practices and benefits*. England, UK: National Foundation for Education Research.
- Hogan, K. (1997). Scaffolding: A powerful toll in social constructivist classrooms. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning. Instructional approaches and issues* (pp. 1–5). Cambridge, MA: Brookline Books.
- Hogan, K., & Pressley, M. (1997). Scaffolding scientific competencies within classroom communities of inquiry. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning. Instructional approaches and issues* (pp. 74–107). Brookline Books: Cambridge, MA.
- Jordan, B., & Putz, P. (2004). Assessment as practice: Notes on measures, tests, and targets. *Human Organization*, 63, 346–358.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.
- Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher*, 28(2), 16–26 46.
- Lampert, M. (2001). *Teaching problems and problems of teaching*. New Haven: Yale University Press.
- Lawson, A. E. (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education*, 25(11), 1387–1408.
- Leahy, S., Lyon, C., Thompson, M., & William, D. (2005). Classroom assessment: Minute by minute, day by day. *Educational Leadership*, 63(3), 18–24.
- Lemke, J. L. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex.

- Lepper, M. R., Drake, M. F., & O'Donnell-Johnson, T. (1997). Scaffolding techniques of expert human tutors. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning. Instructional approaches and issues* (pp. 108–144). Cambridge, MA: Brookline Books.
- Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, 21(3), 335–359.
- Li, M. (2001). *A framework for science achievement and its link to test items*. Stanford, CA: Stanford University. [unpublished dissertation].
- Li, M., Ruiz-Primo, M. A., & Shavelson, R. J. (2006). Towards a science achievement framework: The case of TIMSS 1999. In Howie F. S. & Plomp F. T. (Eds.), *Contexts of learning mathematics and science: Lessons learned from TIMSS* (pp. 291–311). London, UK: Routledge.
- Louca, L. T., Tzialli, D., & Zacharia, Z. C. (2008). Identification–interpretation/evaluation–response: A framework for analyzing classroom-based teacher discourse in science. *International Conference of Learning Sciences* Paper retrieved June 28th 2010, from: <http://portal.acm.org/citation.cfm?id=1599871.1599873>.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Harvard University Press: Cambridge, MA.
- Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, 30(3), 359–377.
- Minstrell, J., & van Zee, E. (2003). Using questioning to assess and foster student thinking. In J. M. Atkin & J. E. Coffey (Eds.), *Everyday assessment in the science classroom* (pp. 61–73). Arlington, VA: National Science Teachers Association Press.
- Minstrell, J., Li, M., & Anderson, R. (2009). *Evaluating science teachers' formative assessment competency*. Technical report submitted to NSF.
- Mortimer, E. F., & Scott, P. H. (2000). Analyzing discourse in the science classroom. In R. Millar, J. Leach, & J. Osborne (Eds.), *Improving science education. The contribution of research* (pp. 126–145). Open University Press: Buckingham, England.
- Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Open University Press: Mainshead, UK.
- Moss, P. A. (2008). Sociocultural implications for the practice of assessment I: Classroom assessment. In P. A. Moss, D. Pullin, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn*. Cambridge University Press: New York.
- Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English*, 25, 261–290.
- Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. (2001). *Questions in time: Investigating the structure and dynamics of unfolding classroom discourse*. The National Research Center on English Learning & Achievement. Report Series 14005. Paper retrieved July 1, from: <http://cela.albany.edu/nystrand01-5/index.html>.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academies Press.
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28, 4–13.
- Roehler, L. R., & Cantlon, D. J. (1997). Scaffolding: A powerful tool in social science constructivist classroom. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning. Instructional approaches and issues* (pp. 6–42). Cambridge, MA: Brookline Books.
- Roth, W. M. (1996). Teacher questioning in an open-inquiry learning environment: Interactions of context, content, and student responses. *Journal of Research in Science Teaching*, 33, 709–736.
- Ruiz-Primo, M. A., & Furtak, E. M. (2004). Informal assessment of students' understanding of scientific inquiry. Paper presented at the meeting of the American Education Research Association.
- Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, 11(3–4), 205–235.
- Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57–84.
- Ruiz-Primo, M. A., Sands, D., & O'Brian, J. O. (2007). *Improving teachers' formative assessment practices and students' learning in mathematics and science classrooms*. Proposal submitted to the Institute of Education Science.
- Sadler, R. D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Sadler, R. D. (1998). Formative assessment: Revisiting the territory. *Assessment in Education*, 5(1), 77–84.
- Scardamalia, M., & Bereiter, C. (1992). Text-based and knowledge-based questioning by children. *Cognition and Instruction*, 9, 177–199.
- Schunk, D. H. (1982). Effects of effort attributional feedback on children's perceived self efficacy and achievement. *Journal of Educational Psychology*, 74(4), 548–556.
- Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18, 337–354.
- Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2007). *Motivation in education. Theory, research, and applications*. Upper Saddle River, New Jersey: Pearson, Merrill Prentice Hall.
- Scott, P. H., Mortimer, E. F., & Aguiar, O. G. (2006). The tension between authoritative and dialogic discourse: A fundamental characteristic of meaning making interactions in high school science lessons. *Cognition and Instruction*, 20, 399–484.
- Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thought, judgments, decisions, and behavior. *Review of Educational Research*, 51(4), 455–498.
- Shute, V. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189.
- Sinclair, J. M., & Coulthard, M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London, UK: Oxford University Press.
- Stipek, D. J. (1984). The development of achievement motivation. In Ames, R., & Ames, C. (Eds.), *Research on motivation in education* (Vol. 1, pp. 145–174). San Diego, CA: Academic Press Inc.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment. Teaching, learning and assessment in the classroom*. Maidenhead, Philadelphia: Open University Press.
- van Zee, E. H., & Minstrell, J. (1997). Using questioning to guide student thinking. *The Journal of the Learning Sciences*, 6, 229–271.
- Vygostsky, L. S. (1978). *Mind in society: The development of higher psychological process*. Cambridge, MA: Harvard University Press.
- Wells, G. (1999). *Dialogic inquiry: Towards a sociocultural practice and theory of education*. Cambridge: Cambridge University Press.
- William, D., & Black, P. (1996). Meaning and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal*, 22(5), 537–548.
- William, D., & Leahy, S. (2007). A theoretical foundation for formative assessment. In J. H. McMillan (Ed.), *Formative classroom assessment: Theory into practice* (pp. 29–42). New York: Teachers College Press.
- Yeany, R. H., & Miller, P. A. (1983). Effects of diagnostic/remedial instruction on science learning: A meta-analysis. *Journal of Research in Science Education*, 20(1), 19–26.