White Paper

Characterizing Observable Formative Assessment Practices:

A Literature-Based Framework

Amelia Wenk Gotwals,
Department of Teacher Education, Michigan State University, East Lansing, USA

Dante Cisterna,
School of Natural Resources, University of Nebraska, Lincoln, USA

John Lane,
Michigan Assessment Consortium, Mason, USA

Dawnmarie Ezzo,
Department of Teacher Education, Michigan State University, East Lansing, USA

Tara Kintz,
Michigan Assessment Consortium, Mason, USA

This work was completed with the support of the Michigan Department of Education for the Formative Assessment for Michigan Educators Program. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agency

Characterizing Observable Formative Assessment Practices: A Literature-Based Framework

Formative assessment involves both teachers and students in the gathering of evidence and use of feedback to modify teaching and learning. This paper synthesizes the literature in three dimensions of observable formative assessment practices based on questions other researchers have identified. The synthesis focuses on characterising observable teaching and learning practices as well as empirical evidence for these practices. These research-backed dimensions can be used to provide a framework that could be used to support teachers in improving their formative assessment practices.

Introduction

Effective use of formative assessment (i.e., Assessment *for* Learning) has long been touted as promoting student learning (e.g., Andersson & Palm, 2017; Black & Wiliam, 1998; Decristan, et al., 2015) and more equitable educational opportunities (e.g., Bishop, Berryman, Cavanagh & Teddy, 2009; Hill, et al., 2017). In response, formative assessment has piqued the interest of researchers, teacher educators, teachers, and testing companies (Bennett, 2011). However, there are nuances in how formative assessment is defined both in research and practice (Dunn & Mulvenon, 2009). In 2011, a critical review of the field found that the current definitions of formative assessment include a broad diversity of approaches with little consistency in what one might expect to see from one implementation of "formative assessment" to another (Bennett, 2011). This paper examines how we might characterize and further articulate observable formative assessment practices that teachers and students use in classrooms. By focusing on what we can see happening in the classroom, we hope to provide a framework for supporting teachers as they develop more sophisticated formative assessment practices.

We draw on the definition of formative assessment developed by the Council of Chief State School Officers (CCSSO) in the United States, which asserts that formative assessment is "... a planned, ongoing process used by all students and teachers during learning and teaching to improve student understanding of intended disciplinary learning outcomes, supporting students becoming more self-directed learners" (CCSSO, 2017, p. 1). The definition provides a holistic vision for formative assessment in the classroom, and it allows for different variations in teaching practices for teachers and learning strategies for students (Heritage, 2007; Shavelson, et al., 2008). The process may include informal, "on-the-fly" interactions between teachers and students or just between students (Cowie & Bell, 2001); it may include discussions or activities where the teacher planned how she will draw out students'

thinking during the course of instruction; and it may include formal tools and activities embedded in the curriculum to garner feedback at key points in the learning process (Cowie & Bell, 2001; Heritage, 2007).

All of these types of formative assessment practices (informal to formal) are important and can improve students' learning (Andersson & Palm, 2017; Black & Wiliam, 1998). Many studies have examined embedded and more formal assessment practices of teachers (e.g., Ayala et. al., 2008; Feldman & Capobanico, 2008; Furtak, et. al., 2008; Shavelson, et. al., 2008). However, few studies have been successful in capturing the nuanced ways in which teachers successfully implement formative assessment practices in the course of instruction. Research (and experience) tells us that classroom instruction is complex and difficult to measure (e.g., Authors, 2018; Jackson, 1990; Rowen & Correnti, 2009). When formative assessment practices are used to support student learning, the line between instruction and assessment is not always clear (Leahy et al. 2005, McNamee & Chen 2005; Wylie & Lyon, 2012); and teaching, learning, and assessing become integrated such that it may be difficult to tease out each component.

In addition, formative assessment practices in the classroom may prove to be problematic to measure or observe, since they are often tacit (Bell & Cowie, 1999, 2001; Mavrommatis, 1997; Wylie & Lyon, 2015). Similar to others (e.g., Wylie & Lyon, 2012; Randel & Clark, 2013), we view formative assessment as a process that is related to, but distinguishable from instruction. Therefore, we posit that characterizing specific dimensions of observable formative assessment practices will allow us to better understand and describe the nuances in teachers' implementation of these practices. With this better understanding, we will be better able to support teachers in the classroom. Below we examine the different ways that formative assessment practices have been parsed in the literature and suggest dimensions of formative assessment that may be visible in the classroom. We then provide a synthesis of research on these potentially visible dimensions of formative assessment.

Characterizing Formative Assessment

Many books and articles have laid out different steps or divisions of formative assessment. For example, Black and Wiliam's (2009) mapping of formative assessment posits that effective formative assessment practice requires the enactment of each of the following steps:

- Teachers and students establish a common understanding of a task's learning goals;
- Teachers elicit student thinking;
- Students respond to teachers' elicitation in ways that reveals their thinking;
- Teachers interpret students' responses to make sense of where students are relative to goals for student learning;
- Teachers take action (e.g., trying a new strategy) to move students in the desired direction, on the basis of their interpretation; and
- Teachers re-assess to measure the action's success.

Heritage (2007) provided a more succinct division of the formative assessment process into four essential elements: (1) identifying the learning gap, (2) feedback, (3) student involvement, and (4) learning progressions. Similarly, Popham (2008) described the key attributes of formative assessment as being: (1) a planned process, (2) assessment-elicited evidence, (3) teachers' instructional adjustments, and (4) students learning tactics adjustments. Whether explicitly stated or not, in each of these definitions teachers and students must play an active role in understanding students' current ideas and moving them to the ultimate goals for learning.

For this paper, we group the essential features of formative assessment listed above into three main dimensions of formative assessment practices that may be observed in the classroom. We utilize work on feedback by Sadler (1989) and Hattie and Timperley (2011)[1] and posit that observable formative assessment practices may be characterized into three dimensions structured around key questions that teachers and students should ask themselves and each other as they move through the learning process.[2]

(1) Use of learning targets and goal setting: Where are we (teachers and students) going?

(2) Evidence of student understanding: What does the student understand now?

(3) Closing the gap/Responding to students: How do we (teacher and students) get to the learning target?

While we have simplified our discussion into three main dimensions, each dimension may include several different observable components. Below we describe each of these dimensions of observable formative assessment practices, the components of each dimension, and examine empirical research that has been done on each of these components (see Table 1).

We used a targeted (or "focused") literature review in order to gather research about each formative assessment dimension and component (Dixon-Woods, Booth, Sutton, 2007; Grant & Booth, 2009; Paré, Trudel, Jaana, & Kitsiou, 2015). A targeted literature review is based on a question or topic and has clear methods for inclusion and exclusion. In addition, targeted literature reviews provide qualitative summaries of important aspects of the researchers' questions. Our questions for the literature review were: (1) how is formative assessment (or assessment for learning) defined in the literature? (2) what are the main teaching and learning practices associated with formative assessment: (3) what does empirical research suggest about the effectiveness of formative assessment teaching and learning? We included both practitioner books as well as peer-reviewed research articles. Our rationale was that the practitioner books would allow us to understand the many different ways formative assessment is

1 Sadler (1989) delineated 3 necessary components of feedback: (1) the standard, which is to be achieved, (2) the actual level of performance, and (3) how to go about closing the gap. Building on this, Hattie and Timperley (2011) suggested that "effective feedback must answer three major questions asked by a teacher and/or a student: Where am I going? (What are the goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?) (p. 86).

2 We believe that formative assessment should be a planned process. However, in this paper, we are focusing on observable classroom practices of teachers, and we do not include the planning process.

defined in the field. The peer-reviewed articles allow us to better understand the findings of empirical studies of formative assessment. While this type of literature is not meant to be comprehensive, it can allow for the synthesis of literature into categories or frameworks (Paré, Trudel, Jaana, & Kitsiou, 2015)

Table 1.

Dimensions and Practices in Formative Assessment

Dimension	Practice	Definition
1: Use of learning targets and goal setting: Where are we (teachers and students) going?	Learning target use	How teacher introduces the learning target to students and connects it to instructional activities
	Learning target: criterion	The extent to which learning targets are rigorous, essential to the discipline, and connected to a clear instructional sequence
	Learning target: conditions	How the teacher explains how students are expected to demonstrate mastery
2: Eliciting evidence of student understanding: What does the student understand now?	Type of question/ information elicited	The cognitive demand of questions and tasks that teachers use to gather evidence of student understanding
	Elicitation Strategies	Strategies to guide questioning and elicit evidence from students
	Self-Assessment	Opportunities for students to be examine their own work and regulate their own learning
3: Closing the gap/ Responding to students: How do we (teacher and	Feedback	The types of feedback provided to students and the manner in which the teacher provides the feedback
	Peer assessment	Opportunities for peers to examine each other's work (and, potentially, help each other learn)

Dimension	Practice	Definition
students) get to the learning target?	Instructional Decisions	The extent to which teachers are explicit about the rationale for making instructional choices alterations
	Use of evidence	The extent to which teachers makes instructional decisions and provides feedback based on student evidence connected with the learning targets

Dimensions of Formative Assessment

Dimension 1: Use of learning targets and goal setting: Where are we going?

As Wiliam and Black (1998) described, assessment refers, "to all those activities undertaken by teachers—and by their students in assessing themselves—that provide information to be used as feedback to modify teaching and learning activities" (p. 140). Formative assessment involves both teachers and students in the gathering of evidence and use of feedback to modify teaching and learning. These practices require that both teachers and learners know what they are trying to accomplish, the extent to which prior attempts to meet learning goals have been successful, and what steps, if any, both students and teachers must take in order to achieve success. Indeed, the ongoing feedback among students and teachers requires clearly articulated learning targets (also known as learning goals or learning objectives) and thus learning targets play a crucial role in the formative assessment process (Popham, 2008; Randel & Clark, 2013).

Learning targets allow teachers to align instruction, assessment, and feedback toward a common learning outcome, ensuring coherence in the learning process (Wiggins & McTighe, 2005). In addition, clearly articulated learning targets facilitate students' understanding of what they are expected to learn. Without clearly communicated learning targets teachers and students may not have the same vision for what is to be learned, causing confusion for the purpose of specific activities and students' abilities to reflect on their own work.

Learning targets can be defined as "statement[s] of what students are expected to demonstrate as a result of instructional activity" (Randel & Clark, 2013, p. 147). The specifics of a quality learning target can be found both in existing literature and in many of the observational frameworks being developed for teacher evaluation in the United States,[3] and internationally (see Martinez, Taut, & Schaaf, 2016).

[3] Observational protocols (e.g. the Danielson Framework, The Observational Protocol Based on the Art and Science of Teaching, *The Thoughtful Classroom*, and *5 Dimensions of Teaching*) each list learning targets as a key element of

Quality learning targets often include rubrics or other tools that help to make the criteria for meeting various levels of performance explicit (Arter & McTighte, 2001) and are connected to local and national standards (Randel & Clarke, 2013). In addition, they

> represent rigorous and important learning in the discipline. The outcomes are clear, written in the form of student learning, and permit viable methods of assessment. Outcomes reflect several different types of learning and, where appropriate, represent opportunities for both coordination and integration. Outcomes take into account the varying needs of individual students (Danielson, 2011, Purpose Standard 1).

In addition, Marzano (2001) posits that, "broadly defined, goal setting is the process of establishing a direction for learning" (p. 93). He argues that learning targets, or goals, should narrow students' focus without overly circumscribing student interest, which might prevent students from seeing the bigger picture and the more distant and broader curricular aim. Finally, he suggests that students should be encouraged to personalize the learning goals. He cites earlier work by Mager (1962) that suggests that effective objectives describe the performance expectations, conditions under which the learning will occur, and a well-articulated threshold for mastery.

One observable feature of learning targets, then, is quality of construction. However, having quality learning targets or objectives is not sufficient for effective integration of these learning targets into formative assessment. Quality use of learning targets necessitates having both teachers and students' employ the learning targets to guide their instruction and learning in meaningful ways. For example, Huinker and Freckmann (2009) argued that effective teacher implementation of learning goals includes four characteristics. First, prior to instruction, the teacher develops a clear and communicable goal of the concepts that students will learn. Second, teachers translate this goal into student-friendly language. Third, students themselves are able to describe the concepts they will be learning or are currently learning. Finally, teachers can articulate the alignment of their lesson progression with district, state, and national standards for learning.

In addition, if learning targets are to be helpful, they should be referenced often enough for students to make connections to the activities and their intended learning. For students, the regulation of their own learning is based on their understanding and use of learning targets (Allal, 2010). Students and teachers need to understand the learning target in similar ways so that they are both aiming for the same outcome. When the students understand the learning target, they are more likely to be able to engage in quality practices discussed below (e.g., self-assessment, peer assessment, use of feedback).

The empirical research on the value of learning targets in and of themselves, however, is quite slim. This is likely due to the fact that learning targets are embedded in the formative assessment process and are not a stand-alone element of instruction that can be implemented or evaluated in isolation. However, researchers have found that when students are provided with clearly articulated learning targets and

---

effective instruction and discuss the qualities of these targets and how they should be embedded in classroom practice, although each has a slightly different conception of quality learning targets.

criteria for quality work, they are better able to evaluate the quality of their own and their peer's work (Roth, et al., 2011; White & Fredrickson, 1998). In addition, when teachers share learning expectations verbally or in rubrics and involve students in the construction of learning goals, instruction tends to become more student-centered and students are more likely to take responsibility for their own work (Andrade, Du & Want, 2008; Tell, Bodone & Addie, 2000).

Dimension 2: Evidence of Student Understanding (What does the student understand now?)

Done well, formative assessment is a dialogic process that allows students' ideas to guide instruction and learning (e.g., see Crossouard & Pryor, 2012). An essential component, then, is engaging students in dialog, both internal and external, to elicit their understanding. Within this dialog, providing opportunities for students to demonstrate their understanding (i.e., eliciting students' ideas) is, perhaps, the most crucial step in the formative assessment process, because without knowing what students know, teachers cannot provide feedback or make informed instructional decisions and students do not have the information to regulate their learning (Heritage, 2007).

This gathering evidence of student understanding can come in many forms. Griffin (2007) argued that gathering evidence of cognitive and affective learning can happen through examining four observable actions: what people (1) say, (2) write, (3) make, or (4) do. As we are examining observable formative assessment practices in the classroom, we focus mainly on what teachers and students say and do. Thus, in this dimension we focus on two observable teacher practices: (1) eliciting students' ideas and (2) promoting self-assessment.

Eliciting Students' Ideas

Questions, either from the student or the teacher, form the basis of gathering information about students' understanding. These questions may be asked in a whole class discussion, small group work, a one-on-one conversation, or on a worksheet or other piece of written work[4]. The observable forms of questioning in a classroom often happen during "assessment conversations" (Duschl & Gitomer, 1997; Ruiz-Primo & Furtak, 2006) where teachers and students (or students together) gather information about what the student knows and use the information for future pedagogical steps (e.g., Ruiz-Primo, 2011), instead of merely providing the correct or acceptable responses (Ruiz-Primo & Furtak, 2007). There are many criteria for examining the nature of the questioning that occurs. For this practice, we examine two observable features of questioning in the classroom: (1) the types of questions are being posed and (2) the elicitation strategies employed.

[4] Assessment tools and students' written responses are often a critical component of gathering evidence of what students know and can do. We do not focus on an examination of tools themselves in this dimension, as we are interested in real-time observations in the classroom. However, many of the criteria discussed may also apply to tools.

Types of questions

Teachers and students ask many types of questions in the course of instruction with various purposes. High quality questions require students to move beyond merely repeating the teacher's ideas or reciting a definition. Questions in classrooms often range from those with a low cognitive demand to high cognitive demand (Henningson & Stein, 1997; Leach & Scott, 2002). This may also be categorized as the depth of knowledge (DoK) required by a question (Webb, 2005, 2009). Questions with lower cognitive demand typically focus on memorization (recall) or tasks that ask students to complete a series of specific steps or plug numbers into a formula. They also tend to focus on procedures, vocabulary definitions, recalling information or seeking to describe or explain 'what' happened. Questions with lower cognitive demand often do not involve much intellectual work (Henningson & Stein, 1997; Silver & Smith, 2015); however these questions may still challenge students.

Questions with higher cognitive demand typically focus on student sense making. These questions ask that students use information or ideas that they have learned to solve a problem, or to consider something from a new perspective or lens. Higher cognitive demand questions ask a great amount of intellectual work from students and often do not have simple, discrete answers. Rather, questions with higher cognitive demand often require students to provide support or explanation in the form of reasoning, experiences, or rationale (Henningson & Stein, 1997; Silver & Smith, 2015).

Questioning in the classroom often will include a range of questions from low to high cognitive demand. However, only asking questions with low cognitive demand limits the types of evidence teachers and students can use. By asking students questions that are high in the cognitive demand required to answer them, teachers can gather important information about students' understandings. In addition to eliciting information about students' ideas and understandings, using a range of carefully crafted questions – both open-ended and more guiding can also provide teachers with opportunities to support students by offering crucial feedback on their learning relative to a learning target and prompt teachers to make instructional decisions based on evidence of students' ideas (Furtak, et al., 2016).

High-quality questions are also those that prompt students to reveal how and why they are thinking about working through a specific problem or idea. Questions that encourage students to reflect on their own metacognitive processing can provide information for both teachers and students. Teachers can use the information about the approach students may take for a particular problem or idea to inform the types of feedback to give and what next instructional moves to make, while students can use the information to regulate their own learning. Metacognitive knowledge involves a student's awareness of his or her own thinking. This 'awareness' includes understanding one's own capacities, limitations, and difficulties; and knowing when and where to use strategies and skills that they have learned (Pressley & McCormick, 1987). This information is helpful to the teacher in making future instructional decisions during teaching (Marzano, 2006), but it can also be a valuable source of information for students who can use that information to help themselves improve their own learning (Pressley & McCormick, 1987), improving student self-assessment.

When teachers ask students high quality questions (i.e., those defined above), they are better able to understand students' ideas (Carpenter, et al., 1989), are better able to provide feedback targeted to

students' specific ideas, are more likely to use student ideas in their planning (Jones & Krouse, 1988), and are able to make instructional decisions that promote student learning (Forman, et al., 2017; Fuchs, et al., 1991; Peterson, et al., 1989).

Elicitation strategies

There has been substantial research on different strategies that teachers use in eliciting students' ideas (e.g., Chin, 2007; Franke, et al., 2009; van Zee, et al., 2001). In addition to procedural concepts such as increasing wait time to increase student processing and engagement (e.g., Rowe, 1986; Tobin & Capie, 1982), the ways in which teachers ask questions and allow for student responses have implications for formative assessment. For example, the ways teachers structure their assessment conversation can influence the type of evidence of student understanding they may elicit; and the ways that teachers follow-up these initial questions have implications for the types of information that students may be willing to share (Davis, 1997) and for the larger classroom culture (Supovitz & Turner, 2000). The most effective questioning strategies privilege and follow students' ideas, as opposed to focusing on getting the 'right answer' (e.g., diSessa & Minstrell, 1998; Otero, 2006).

One of the most documented questioning strategies is the I-R-E format (i.e., teacher initiates (I) dialogue by asking a question, a student offers a response (R) to the question, and then the teacher evaluates (E) the student's response (Mehan, 1985; Cazden, 2001; Reznitskaya, 2012). Similar to this strategy, Kennedy (2005) identified 'guessing games,' where the teacher poses questions and rejects students' ideas until she gets the idea she wants or dismisses student ideas that do not conform to the desired response, and 'funneling questions,' when the teacher asks a series of questions that guide the students through a procedure or to a desired end (Herbel-Eisenmann & Breyfogle, 2005; p.485). In all these situations (I-R-E, guessing games, funneling questions), the teacher is the one engaged in cognitive activity and the student is merely answering the questions to arrive at an answer, often without seeing the connection among the questions.

As teachers begin to probe for student understanding, students' ideas can become more central to the assessment conversation. For example, teachers may gain clearer understandings of student thinking by following a student's response to their initial question with questions that probe or press students for additional information. Kennedy (2005) called these questions 'I-R-probes,' which begin similarly to an IRE questioning pattern, but instead of evaluating a student response, teachers may ask for more information. Van Zee and Minstrell (1997) called this a 'reflective toss,' where the teacher puts some of the responsibility of meaning-making back on the student, by asking questions the help students to clarify their meanings, to consider additional viewpoints, and to monitor their own thinking.

In addition to gaining insight into student thinking by using questions to support students in clarifying and extending their ideas, significant insight into student understanding is gained when teachers blend instructional and eliciting strategies by working alongside their students in 'joint acts of meaning-making and knowledge construction' (Wolfe & Alexander, 2008; p. 1). As students and teachers consult and work together to develop the students' understanding and progress on a particular task or learning goal, teachers can simultaneously scaffold students' developing understandings as they work to elicit student thinking and sense-making (Ruiz-Primo & Furtak, 2006). Teachers who work with their students and ask questions that are intended to both elicit and scaffold student thinking gain richer insight into student

thinking, enabling them to 'interpret student responses in the moment and determine next pedagogical steps, including feedback, that will move the students closer to successful performance' (Heritage & Heritage, 2013, p. 188).

In characterizing these types of assessment conversations, Ruiz Primo and Furtak (2006) suggested the use of cycles named ESRU in which the teacher: *E*licits students' ideas; *S*tudents respond; the teacher *R*ecognizes students´ responses; and *U*ses the information to promote future learning. Throughout these conversations, teachers can scaffold understanding and build communities in which students begin to take more ownership for their learning (Forman, Ramirez-DelToro, Brown & Passmore, 2017). Students in classrooms with teachers who scaffold these type rich assessment conversations tend to score higher than those in classrooms with direct instruction (O'Connor, et al., 2015), with both vocal and less-vocal students benefitting from this dialog-rich classroom context (O'Connor, Michaels, Chapin & Harbaugh, 2017).

In order to begin dialog, teachers may elicit student responses using tools and strategies such as each student having a white board or clicker to record their response to a question (e.g., Lee, Feldman, & Beatty, 2012) or by asking students to model a phenomenon (Windschitl, Thompson, Braaten & Stroupe, 2012). Teachers can then use this information to guide the types of questions that they ask and how they probe for student understanding. Similarly, teachers may use exit tickets to inform them of where students' understanding is at the conclusion of a lesson and use this information to inform their future instruction.

Self-assessment

Student self-assessment has long been recognized as a key component of the formative assessment process (Andrade, Du, & Wang, 2008; Black & Wiliam, 1998). Black and Wiliam (1998) characterized the link between formative assessment and self-assessment as 'inevitable' and argue that, 'if formative assessment is to be productive, pupils should be trained in self-assessment so that they can understand the main purposes of their learning and thereby grasp what they need to achieve' (p. 7).

Drawing a distinction between self-evaluation (e.g., concepts of self-efficacy), Brown and Harris (2013) contend that self-assessment is, 'about judging, evaluating, and considering one's own academic work or abilities' (p. 369). Andrade and Valtcheva (2009) provided a more detailed definition. They describe self-assessment as, 'a process of formative assessment during which students reflect on the quality of their work, judge the degree to which it reflects explicitly stated goals or criteria, and revise accordingly' (p. 13). Heritage (2013) states that self-assessment has two aspects: (1) self-appraisal and (2) self-management (Paris & Winograd, 1990). Self-appraisal includes students' ability to be metacognitive about their current state of learning, while self-management includes students' abilities to take the next steps in their learning.

There are myriad ways that teachers and students can enact self-assessment as defined above and many scholars have detailed steps that teachers and students can take to this end. For example, Andrade and Valtcheva (2009) provided a useful delineation of necessary qualities of successful self-assessment practices. First, teachers must clearly articulate expectations for student performance (i.e., learning targets and criteria for success). Second, teachers must allow for students to practice self-

assessment before the conclusion of an assignment or project. In other words, teachers and students should use self-assessment as a tool to propel student learning and achievement before summative evaluation. Finally, teachers should require students to use the understanding gained from their self-assessments to revise their work and move toward the learning target.

Despite its intuitive appeal and nearly universal scholarly support (Brown & Harris, 2013), classroom teachers employ student self-assessment infrequently as a method for improving student learning (Hunter, Mayenga, & Gambell, 2006). Brown and Harris (2013) speculate that this reluctance stems from 'tensions teachers report between the use of student-led assessment practices and the externally and teacher-controlled summative results generally reported to stakeholders' (p. 367).

Engaging students in the process of self-assessment has a significant impact on student learning and performance. Several studies focusing on self-assessment found that students using self-assessment strategies experienced a faster rate of learning than students who did not use self-assessment strategies (e.g., Andrade & Brookhart, 2016; Andrade, Du & Wang, 2008; Fontana & Fernandes, 1994; Graham, Hebert & Harris, 2015; White & Frederiksen, 1998). In their review of self-assessment research, Brown and Harris (2013) conclude, 'there is empirical evidence that self-assessment of a task or self-confidence in the quality of the work will generally improve academic performance across a range of grade levels and subject areas' (p.381). Similarly, Fuchs et al. (2000) argue that providing students with standards for quality and teaching them to apply those standards through scoring their work with a rubric can not only dramatically increase student performance, but can also enable students to apply these standards to future work.

Dimension 3: Responding to Students' Ideas/Closing the Gap (How do we [teacher and students] get to the learning target?)

The evidence that teachers and students gather that pinpoint the current status of a student's learning should be used to inform the next steps in the learning process. As Heritage (2013) states, '…for formative assessment to be valid, it must yield tractable insights' (Heritage, 2010, p. 185). These insights into students' understanding can be used to guide subsequent teacher and student moves. These subsequent moves often take the form of feedback (either from the teacher to student or between peers) or the next instructional or learning decision. The goal of gathering evidence of student understanding is to use it to guide activities, discussion, or other methods to close any gap between students' current understanding and the learning target. In addition, the intention is to support student agency to close the gaps in their own learning.

Feedback

Feedback is information that flows to and from the teacher and students or between students and is conceptualized as 'information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding…' (Hattie & Timperley, 2007; p. 81) and this information is 'designed to reduce discrepancies between current understandings or performance and a desired goal' (Wiliam, 2013). The goal of feedback is to provide clear and descriptive information to help students close the gap between their current understanding and the learning goal (Sadler, 1989).

Sadler's (1989) model emphasizes feedback to students through the use of a 'feedback loop.' This loop involves both teachers and students in an ongoing process. Feedback, both verbal and written, can support students by showing them how close they are to the learning targets and what they can do to move closer (Landauer, Lockbaum & Dooley, 2009).

There has been a plethora of research on what constitutes effective feedback (e.g., Brookhart, 2008; Hattie & Temperley, 2007; McManus, 2008; Nicol & Macfarlane-Dick, 2005; Wiggins, 2016). There is much agreement that feedback works to support student learning when it is timely, ongoing, actionable, and consistent. In addition, rather than being evaluative, feedback should focus on aspects of the student's work rather than on the student as a learner and should be descriptive of that work, revealing to the student how to do better the next time. Effective feedback from teachers also provides clear, descriptive, criterion-based information that indicates to the students where they are relative to a learning goal, how their understanding differs from the desired learning goal, and how they can move forward.

In their meta-analysis, Hattie and Temperley (2007) found that when feedback was clear and related to learning targets, it had a large effect on student learning. They determined that the most effective place for assessment-related feedback is during processing (i.e., when students work through strategies for doing/completing problems) and that feedback can support students in improving their self-assessment and metacognitive abilities. Studies showing the highest effect sizes involved students receiving informational feedback about a task and how to do it more effectively. Lower effect sizes were related to praise, rewards, and punishment (Hattie & Temperley, 2007).

Parr and Timperley (2010) defined quality feedback based on three criteria: 1) where students were positioned in relation to the performance desired, 2) information on the desired performance, and 3) what students needed to accomplish to achieve the desired performance. Other research indicates that feedback is most effective when it is understood by the student, addresses the mistakes the student made, and prompts students rather than provides an answer (Sally, 2012). Feedback can also take the form of questions. Feedback questions (e.g., what evidence can you provide to support that idea?) can support students in examining their own learning and success, or to deepen their own understanding relative to a concept or idea (Chin, 2006).

The type of observable feedback seen in the classroom might be described as a feedback loop that is part of a larger questioning cycle. However, most of the research is about written feedback. In addition, despite the ubiquitous support for the importance of feedback, much of the

> …knowledge base about feedback is … drawn from studies done in laboratories or in artificial classroom environments…. Furthermore, we are aware of many strategies and techniques suggested in books and papers… however, to some degree, it can be said that they have not been empirically tested (Ruiz-Primo & Li, 2013, p. 215-216).

This may be because, as opposed to written feedback on work, observable feedback usually happens quickly; and thus, may be difficult to observe.

Peer assessment

Topping (2009) defines peer assessment as 'an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal status learners' (p. 20). Used in different task and contexts, peer-assessment provides opportunities to students to reflect about their learning (CCSSO, 2008). Peer-assessment allows students to practice a different classroom role, by examining others' work, and to interpret learning targets and assessment criteria in students' own and everyday language. Peer assessment holds potential for engaging students with different types of feedback; for example, students tend to perceive feedback from peers and teachers differently (Topping, 2009). In addition, students' feedback can often be more immediate than teachers', may be perceived as more open to negotiation and less authoritative (Cole, 1991), and provide the opportunity to examine differences of opinion as opportunities for learning (Van Gennip, Segers, & Tillema, 2010).

Peer assessment is a practice that needs to be supported by teachers, especially to help students be reflective on peers' work and provide descriptive feedback. Implementation of peer assessment in the classroom has been used to help students think about criteria for success (Webb & Jones, 2009). Thus, the use of rubrics that specify these criteria are important for setting clear expectations of students' work and providing feedback to peers accordingly (Sato & Atkin, 2007; Wiliam, 2006).

As a process, the use of peer assessment has the potential to provide more individualized and detailed feedback to students compared with the feedback only provided by the teacher (Topping, 2009). It also provides opportunities for students to contrast peers' ideas and work based on evidence (Sadler, 1989) and suggest pathways for improvement and moving to higher levels of performance (Wiliam, 2006). However, the enactment of peer assessment is a time-consuming process that requires continuous teacher scaffolding (Gioka, 2009; Webb & Jones, 2009). Some teachers have observed that students find examining others' work easier than examining their own (Wiliam, 2006). Thus, peer assessment may be a first step in scaffolding student self-assessment (Sliwka & Spencer, 2005; Webb & Jones, 2009). Activating learners as instructional resources for one another – getting students involved with each other in discussions and working in groups – can help improve student learning (King, 1992; Wiliam, 2009).

Instructional decisions

Most definitions of formative assessment include an instructional improvement component of verifying or modifying teaching and learning activities based on evidence of students' understanding as it relates to learning targets (e.g., see Wiliam, 2009). If students have not yet reached learning targets, teachers may make a change in: (1) lesson plans, (2) instructional approaches, or (3) types of assessments to better understand what students know or do not understand. However, the types of practices that are observable in a given class period may be limited. These moments often happen on the fly. Wiliam and Leahy (2007) categorized these as 'moments of contingency' where teachers are continually considering evidence to guide their instructions. Ruiz-Primo and Furtak (2006) found that students whose teachers used the information collected in assessment conversations to push student learning forward tended to have better scores on summative assessments. Specific examples of observable instructional decisions include: re-teaching or teaching differently (e.g., using different examples or wording); grouping students differently; providing different activities; or adjusting assessments to better assess learning. Black and Wiliam (2009) discussed the process of making instructional decisions as having two steps: (1)

diagnostic, i.e., interpreting students' responses in terms of what it reveals about the students' thinking and motivation, and (2) prognostic, choosing an optimum response. Both steps involve complex decisions and often these decisions are made instantaneously.

While there are some guiding principles for instructional decisions, an instructional decision cannot be interpreted outside the context in which it occurred (i.e., it is contingent on both what precedes the decision in terms of both the content and students involved). In addition, an instructional decision that is considered effective in one subject or in a lesson that was based on a particular learning aim within that subject, might be inappropriate for another subject or for an interaction that had a different learning goal in mind (Black & Wiliam, 2009). Despite these limitations, research suggests that instructional decisions should be based on meeting students' learning needs, and these decisions should be made such that the next steps in instruction are likely to be better or better founded than the decisions they would have taken in the absence of the evidence that was elicited (William, 2009).

Sadler (1989) points out the importance of helping students learn to use assessment evidence and modify their learning processes. Including students in the learning process is a key aspect of formative assessment and teachers need to transfer some of the responsibility for making evaluative decisions from teacher to learner. There are some research findings that have focused on teachers' instructional adjustments and decisions based on students' data. Fuchs et al. (1991) found that teachers who systematically monitor student learning, engage with an expert system to plan instructional adaptation, and make appropriate adjustments produce greater gains in student achievement than comparable teachers. Moreover, teachers' use of detailed information about students' skill level to provide meaningful and customized activities helped students to improve math, science, and reading achievement (Bergan et al., 1991). Although making instructional decisions is crucial to provide meaning to the use of formative assessment, its identification and characterization pose challenges for teachers (e.g., Oláh, Lawrence, & Riggan, 2010) as well as classroom observation. Instructional decisions that teachers and students make are complex processes that require using and weighing different types of assessment evidence that may go beyond the observational context. For example, how teachers interpret and understand students' ideas is critical to providing appropriate feedback and making instructional decisions (Furtak, et al., 2016). However, these interpretations and decisions are not usually observable and often involve multiple factors (e.g., student understanding, timing issues, resources).

Given the tacit nature of instructional decisions, teacher interviews may be necessary as part of a classroom observation to make some of the more tacit aspects of practice explicit (e.g., Banhart & van Es, 2015). An interview provides the opportunity for a teacher to explain the process of gathering and interpreting student evidence to inform future instructional decisions, which are difficult to observe. Teachers can explain changes they made in lesson plans, instructional approaches, or types of assessments based on evidence of students' understandings. A simulated recall interview in which teachers observe videos of themselves teaching may facilitate teachers' ability to distinguish and recall "on the fly" instances or "moments of contingency" in which teachers spontaneously and continuously consider evidence to inform their instruction. These interviews would provide the opportunity for teachers to further explain the complex nature of adjusting instruction based on assessment

information, which can be one way of supporting teachers in being reflective about their formative assessment practices as well as being a tool to facilitate the further development of formative assessment research and practice.

Conclusion

Formative assessment has been shown to improve student learning (Black & Wiliam, 2009). While some analyses have called into question the magnitude of the effect size on student learning (e.g., Bennett, 2011; Briggs, et al., 2012; Kingston & Nash, 2011), there is still consensus that, when done well, formative assessment is a critical practice that can increase student engagement, promote student ownership and self-regulation of learning, and can improve student achievement.

In this paper, we identified three dimensions of formative assessment, defined the observable component practices in each dimension, and summarized empirical research on these practices. We understand that these dimensions do not follow a linear path or represent steps in a process. Rather we suggest that these dimensions may be the beginnings of a framework that can be used to draw teachers' attention to salient formative assessment practices and to support teachers as they reflect on their teaching.

Characterizing observable component practices in each dimension blurs the line between good teaching and formative assessment. In addition, many of these component practices are contingent on each other (e.g., clearly defined and communicated learning targets guide the types of questions and feedback that would be effective in promoting student learning). However, perhaps because significant content and pedagogical expertise is required for high-quality implementation of formative assessment (Bennett, 2011), implementation of high-quality formative assessment practices difficult to find, even when teachers have participated in professional development and have access to resources (e.g., Authors, 2015). Therefore, having a framework that illustrates the ways in which formative assessment is an overall instructional process with three large dimensions (i.e., the questions), as well as articulating specific practices within each dimension, can support teachers in understanding ways in which formative assessment practices can be integrated into their instruction.

In our research (i.e., Authors, 2015; 2018) we found that many teachers struggle to make explicit connections between components of formative assessment, for example, they may use learning targets at the beginning of a lesson or instructional unit, but not refer back to the learning targets to guide the collection of student evidence and feedback. This type of framework may be able to support teachers in understanding what these dimensions may look like and the ways they connect with one another.

Articulating the dimensions and component practices of formative assessment may be helpful for supporting teachers' learning of these practices. By clarifying the aspects of formative assessment practices that may be observable in the classroom, we can provide guidance for the development of tools that can support teachers in their professional growth. For example, we can develop observational measures that can allow teachers to reflect on their formative assessment practices, either alone or with a coach or colleague. Since these dimensions of formative assessment are applicable for different units of instructional time (lessons, units), they may provide another layer to support teachers as they plan for and enact instruction that consistently prioritizes students' ideas and learning. Teachers likely already

have resources and enact practices that map onto a dimension of formative assessment. By providing a coherent framework, we can support teachers in aligning their current practices such that when students consider three questions (i.e., Where am I going? Where am I now? How to do I get to the learning target?) they have a clear path to success.

References

Allal, L. (2010), Assessment and the Regulation of Learning. In: P. Peterson, E. Baker, & B. McGaw, (Eds), International Encyclopedia of Education. volume 3, pp. 348-352. Oxford: Elsevier.

Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, *49*, 92-102.

Andrade, H., & Brookhart, S. M. (2016). The Role of Classroom Assessment in Supporting Self-Regulated Learning. In *Assessment for Learning: Meeting the Challenge of Implementation* (pp. 293-309). Springer International Publishing.

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice*, *27*(2), 3-13.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12-19.

Arter, J. & McTighe, J. (2001). Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance. Thousand Oaks, CA: Corwin Press.

Authors (2018). Details omitted for peer review.

Authors (2015). Details omitted for peer review.

Ayala, C. C., Shavelson, R. J., Ruiz-Primo, M.A., Brandon, P. R., Yin, Y., Furtak, E., et al. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education*, *21*(4), 315-334.

Banhart, T. & van Es, E. (2015). Studying teacher noticing: Examining the relationship among pre-service science teachers; ability to attend, analyze and respond to student thinking. *Teaching and Teacher Education*. 45, 83-93.

Bell, B. & Cowie, B. (2001). The characteristics of formative assessment in science education. Science Education. 85: 536-553.

Bennett, R. E. (2011). Formative assessment. A critical review. *Assessment in Education: Principles, Policy & Practices*, 18(1), 5–25.

Bishop, R., Berryman, M., Cavanagh, T., & Teddy, L. (2009). Te Kotahitanga: Addressing educational disparities facing Māori students in New Zealand. *Teaching and Teacher Education*, 25, 734–742.

Black, P., & Wiliam, D. (1998). Inside the black box. *Phi Delta Kappan, 80*(2), 139-148.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, & Accountability, 21*(1), 5-31.

Briggs, D. C., Ruiz-Primo, M.A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, 31(4), 13-17.

Brookhart, S.M. (2008). How to give effective feedback to your students. Alexandria, VA.: ASCD Press

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE.

Carpenter, T.P., Fennema, E., Peterson, P.L., Chiang, C., & Loef, M. (1989). Using knowledge of children's mathematical thinking in classroom teaching: An empirical study. *American Educational Research Journal.* 26(4), 499-531.

CCSSO, (2008). *Attributes of effective formative assessment*. Washington, D.C., Council of Chief State School Officers.

Chin, C. (2006). Classroom interaction in science: Teacher questioning and feedback to students' responses. International Journal of Science Education, 28, 1315–1346.

Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of research in Science Teaching*. 44, 815-843

Cowie, B., & Bell, B. (2001). A model of formative assessment in science education. *Assessment in Education*, *6*(1), 101-116.

Cole, D. A. (1991). Change in self-perceived competence as a function of peer and teacher evaluation. Developmental Psychology, 27, 682–688.

Crossouard, B. & Pryor, J. (2012). How Theory Matters: Formative Assessment Theory and Practices and Their Different Relations to Education. *Studies in Philosophy and Education*. 31, 3, 251-263.

Danielson, C. (2011). The framework for teaching evaluation instrument, 2011 edition. Princeton, NJ, The Danielson Group.

Davis, B. (1997). Listening for differences: An evolving conception of mathematics teaching. Journal for Research in Mathematics Education, 28(3), 355–376.

Decristan, J., Klieme, E., Kunter, M., Hochweber, J., Büttner, G., Fauth, B., ... & Hardy, I. (2015). Embedded Formative Assessment and Classroom Process Quality How Do They Interact in Promoting Science Understanding?. *American Educational Research Journal*, 53(6). 1133-1159.

di Sessa, A. A., & Minstrell, J. (1998). Cultivating conceptual change with benchmark lessons. In J. G. Greeno & S. V. Goldman (Eds.), *Thinking practices in mathematics and science learning* (pp. 155-187).

Dixon-Woods, M., Booth, A., & Sutton, A.J. (2007) Synthesising qualitative research: a review of published reports. *Qualitative Research, 7*, 375–422.

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.

Feldman, A. & Capobianco, B.M. (2008). Teacher learning of technology enhanced formative assessment. Journal of Science Education and Technology. 17(1): 82-99.

Fontana, D., & Fernandes, M. (1994). Improvements in the mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology*, 64, 407–417.

Forman, E.A., Ramirez-DelToro, V., Brown, L & Passmore, C. (2017). Discursive strategies that foster an epistemic community for argumentation in a biology classroom. *Learning and Instruction*. *48*, 32-39.

Franke, M.L., Webb, N.M., Chan, A.G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education. 60*, 380-392.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P.M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, 28(3), 617–641.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., Dutka, S., & Katzaroff, M. (2000). The importance of providing background information on the structure and scoring of performance assessments. *Applied Measurement in Education*,*13*(1), 1-34.

Furtak, E. M., Ruiz-Primo, M. A, Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R., & Yin, Y. (2008). On the Fidelity of Implementing Embedded Formative Assessments and Its Relation to Student Learning. *Applied Measurement in Education*, *21*(4), 360–389.

Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de Leon, V., Morrison, D. & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: Findings from a four-year intervention study. *Instructional Science*. 44, 267-291.

Ginsburg, H. P. (2009). The challenge of formative assessment in mathematics education: Children's minds, teachers' minds. *Human Development*, *52*(2), 109-128.

Gioka, O. (2009) Teacher or examiner? The tensions between formative and summative assessment in the case of science coursework. *Research in Science Education, 39*, 411-428.

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, *115*(4), 523-547.

Grant, M.J. & Booth, A. (2009) A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26, pp. 91–108.

Griffin, P. (2007). The comfort of competence and the uncertainty of assessment. *Studies in Educational Evaluation*, 33, 87–99.

Hattie, J. & Timperley, H. (2011). *The power of feedback*. Review of Educational Research. 77(1): 81-112.

Henningson, M. & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom based factors that support and inhibit mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524 - 549.

Herbel-Eisenmann, B. & Breyfogle, M. (2005). Questioning our patterns of questioning. *Mathematics Teaching in the Middle School. 10(9),* 484-489

Heritage, M. (2013). Gathering evidence of student understanding. In J.H. McMillan (Ed) Handbook of Research on Classroom Assessment. Sage: Thousand Oaks, CA.

Heritage, M., & Heritage, J. (2013). Teacher Questioning: The Epicenter of Instruction and Assessment. *Applied Measurement in Education*, *26*(3), 176-190.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan, 89*(2), 140–146.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24–31.

Hill, M.F., Ell, F., Grudnoff, L., Haigh, M., Cochran-Smith, M., Chang, W-C & Ludlow, L. (2017). Assessment for equity: Learning how to use evidence to scaffold learning and improve teaching. *Assessment in Education: Principles, Policy & Practice* 24(2) 185-204.

Huhta, Ari (2010). *Diagnostic and Formative Assessment.* In Spolsky, Bernard and Hult, Francis M. The Handbook of Educational Linguistics. Oxford, UK: Blackwell. pp. 469–482.

Huinker, D., & Freckmann, J. (2009). *Linking principles of formative assessment to*

classroom practice. Wisconsin Teacher of Mathematics, 60(2), 6-11.

Hunter, D., Mayenga, C., & Gambell, T. (2006). Classroom assessment tools and uses: Canadian English teachers' practices for writing. *Assessing Writing, 11*(1), 42-65.

Jackson, P. W. (1990). Life in classrooms. New York: Teachers College Press.

Jones, E. D., & Krouse, J. P. (1988). The effectiveness of data-based instruction by student teachers in classrooms for pupils with mild learning handicaps. Teacher Education and Special Education , 1(1), 9–19.

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist*, 27(1), 111-126.

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. Educational Measurement: Issues and Practice, 30(4), 28–37.

Landauer, T.K., Lochbaum, K.E., & Dooley, D. (2009). A new formative assessment technology for reading and writing. *Theory Into Practice* 418(1), 44-52.

Leach, J. & Scott, P. (2002). Designing and evaluating science teaching sequences: An approach to drawing upon the concept of learning demands and a social constructivist perspective on learning. Studies in Science Education. 38: 115-142.

Leahy, S., Lyon, C., Thompson, M., & Wiliam D. (2005). Classroom Assessment - Minute by Minute, Day by Day. *Educational Leadership, 63*, 18-24.

Lee, H., Feldman, A. & Beatty, I.A. (2012). Factors that affect science and mathematics teachers' initial implementation of technology-enhanced formative assessment using a classroom response system. *Journal of Science Education and Technology*. 21, 523-539.

Mager, R. (1962). *Preparing instructional objectives*. Palo Alto: Fearon Publishers.

Martinez, F., Taut, S., & Schaaf, K. (2016). Classroom observation for evaluating and improving teaching: An international perspective. *Studies in Educational Evaluation*, 49, 15-29.

Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria: Association for Supervision and Curriculum Development.

Marzano, R. J. (2006). *Classroom assessments and grading that work.* Alexandria, VA: Association for Supervision and Curriculum Development.

Mavrommatis, Y. (1997). Understanding assessment in the classroom: Phases of the assessment process – the assessment episode. Assessment in Education 4(3): 381-400.

McCormick, C. B., & Pressley, M. (1997). *Educational psychology: Learning, instruction, assessment*. New York: Longman.

McManus, S. (2008). *Attributes of effective formative assessment. CCSSO: FAST SCASS.*

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, Virginia: NSTA Press.

Nicol, D., Macfarlane-Dick, D. (2005). *Rethinking Formative Assessment in HE: a theoretical model and seven principles of good feedback practice*. Quality Assurance Agency for Higher Education.

O'Connor, C., Michaels, S., & Chapin, S. (2015). "Scaling down" to explore the role of talk in learning: from district intervention to controlled classroom study. In L.

Resnick, C. Asterhan, & S. Clarke (Eds.), *Socializing Intelligence Through Academic Talk and Dialogue* (pp. 111-126). Washington D.C.: American Educational Research Association.

O'Connor, C., Michaels, S., Chapin, S. & Harbaugh, A.G. (2017). The silent and the vocal: Participation and learning in whole class discussion. *Learning and Instruction* 48, 5-13.

Oláh, L. N., Lawrence, N. R., & Riggan, M. (2010). Learning to learn from benchmark assessment data: How teachers analyze results. Peabody Journal of Education, 85, 226–245.

Otero, V. (2006). Moving beyond the "Get it or don't" conception of formative assessment. *Journal of Teacher Education*, 57(3), 247–255.

Paré, G., Trudel, M., Jaana, M., Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52, pp. 183–199.

Parr, J. & Timperley, H. (2010). Feedback to writing, assessment for teaching and learning and student progress. *Assessing Writing, 15*, 2. 68-85.

Peterson, P.L., Carpenter, T., & Fennema, E. (1989). Teachers' knowledge of students' knowledge in mathematics problem solving: Correlational and case analysis. *Journal of Educational Psychology,* 81(4), 558-569.

Popham, W. J. (2008). *Transformative assessment*. Alexandria: Association for Supervision and Curriculum Development.

Pressley, M., & McCormick, C. B. (1987). *Advanced educational psychology for educators, researchers, and policy makers.* New York: HarperCollins.

Randel, B. & Clark, T. (2013). Measuring classroom assessment practices. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE.

Reznitskaya, A. (2012). Dialogic teaching: Rethinking language use during literature discussions. The Reading Teacher, 65(7), 446–456. doi:10.1002/TRTR.01066.

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K. and Wickler, N. I.Z. (2011), Videobased lesson analysis: Effective science PD for teacher and student learning. J. Res. Sci. Teach., 48: 117–148. doi: 10.1002/tea.20408.

Rowe, M.B. (1986). Wait time: Slowing down may be a way of speeding up! *Journal of Teacher Education* 37, 43-50.

Rowen, B. & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from the study of instructional improvement. Educational Researcher 38(2): 120-131.

Ruiz-Primo, M.A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students for science learning. Special issue in assessment *for* learning, Studies of Educational Evaluation, 37(1), 15-24

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. & Li, M. (2013). Examining formative feedback in the classroom context: New research perspectives. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment*. Los Angeles: SAGE.

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

Sally, J. (2012). Student engagement with assessment and feedback: Some lessons from short-answer free-text E-assessment questions. *Computers & Education, 58,* 2. 818-834.

Sato, M. & Atkin, J. M. (2006). Supporting change in classroom assessment. *Educational Leadership*, 64(4), 76-79.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M., . . . Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. Applied Measurement in Education, 21(4), 295–314.

Shavelson, R. J., Yin, Y., Furtak, E.M., Ruiz-Primo, M.A., Ayala, C.C. … Pottenger, F.M. (2008). In J. Coffey, R. Douglas & C. Steams (Eds.) *Assessing science learning: Perspectives from research and practices.* National Science Teacher Association: Washington D.C. (p. 21-36).

Schneider, M. C., & Randel, B. (2009). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade & G. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 251 -276). New York, NY: Routledge.

Silver, E. A., & Smith, M. S. (2015). Integrating powerful practices: Formative assessment and cognitively demanding mathematics tasks. *Annual perspectives in mathematics education: Assessment to enhance teaching and learning*, 5-14.

Sliwka & Spencer, (2005). , "Scotland: Developing a Coherent System of Assessment", In: *Formative assessment – improving learning in secondary classrooms*. (pp. 205-220). Centre for Educational Research and Innovation. Paris: OECD Publishing. ISBN: 92-64-00739-3.

Supowitz, J.A. & Turner, H.M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching*. 37(9): 963-980.

Tell, C.A., Bodone, F.M., & Addie, K.L. (2000). A framework of teacher knowledge and skills necessary in a standards-based system: Lessons from high school and university faculty. *Paper presented at the annual meeting of the American Educational Research Association,* New Orleans, LA.

Tobin, K. G., & Capie, W. (1982). Relationships between classroom process variables and middleschool science achievement. Journal of Educational Psychology , 74(3), 441–454.

Topping, K. (2009). Peer Assessment. *Theory Into Practice*, 48, 20-27.

Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, *20*(4), 339–343. doi:10.1016/j.learninstruc.2009.08.003

Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: the role of interpersonal factors and conceptions. Learning and Instruction, 20(4), 280e290.

van Zee, E. H., Iwasyk, M., Kurose, A., Simpson, D., & Wild, J. (2001). Student and teacher questioning during conversations about science. *Journal of Research in Science Teaching, 38*(2), 159-190.

van Zee, E. & Minstrell, J. (1997). Using questioning to guide student thinking. The Journal of the Learning Sciences. 6(2) 227-269.

Webb, M., & Jones, J., (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *16*(2), 165-184.

Webb, Norman L. and others. "Web Alignment Tool" 24 July 2005. Wisconsin Center of Educational Research. University of Wisconsin-Madison. 2 Feb. 2006. <http://www.wcer.wisc.edu/WAT/index.aspx>.

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

Wiggins, G. (2016). Seven keys to effective feedback. *On Formative Assessment: Readings from Educational Leadership (EL Essentials)*, 24.

Wiggins, G. & McTighe, J. (2005). *Understanding by Design*. Alexandria, VA: Association for Supervision and Curriculum Development.

Wiliam, D. (2013). Feedback and instructional correctives. In J.H. McMillan (Ed) Handbook of Research on Classroom Assessment. Sage: Thousand Oaks, CA.

Wiliam, D. (2006). Assessment: Learning communities can use it to engineer a bridge connecting teaching and learning. *Journal of Staff Development, 27*(1), 16-20.

Wiliam, D. (2009). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. Cizek (Eds.), *Handbook of Formative Assessment* (pp. 18-40). New York, NY: Routledge.

Windschitl, M., Thompson, J., Braaten, M & Stroupe, D. (2012). Proposing a core set of instructional practices and tools for teachers of science. *Science Education*. 96, 878-903.

Wolfe, S., & Alexander, R. J. (2008, December). *Argumentation and dialogic teaching: Alternative pedagogies for a changing worl*d. London, UK: FutureLab (Beyond Current Horizons).

Wylie, E. C. & Lyon, C.J. (2012). Quality instruction and quality formative assessment: The same or different? Paper presented at the meeting of the American Educational Research Association, Vancouver, B.C.

Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 140-160.