Moving Forward with Caution:

Assessment Research and Practice

Brian Preston

Informing the practice of teaching using formative and interim assessment: A systems approach is a compilation of papers presented at the 11th conference of the Maryland Assessment Research Center for Education Success (MARCES). As one who provided professional development programs on formative assessment to K-12 teachers and administrators for several years, and who looked to several of these

authors for background and inspiration for my own work, the title spoke to me, suggesting it would offer concrete suggestions of immediate benefit in classrooms. Some of the chapters are more theoretical than practical, but all offer something for those building backgrounds in balanced assessment or developing coursework in formative and interim assessment for preservice or inservice training programs. Most authors provide useful introductions to the background research, offer definitions and define terms, cite key researchers, and support the efficacy of formative and interim assessments, building a strong and consistent position in favor of increasing the use of such assessments as a tool for improving student achievement.
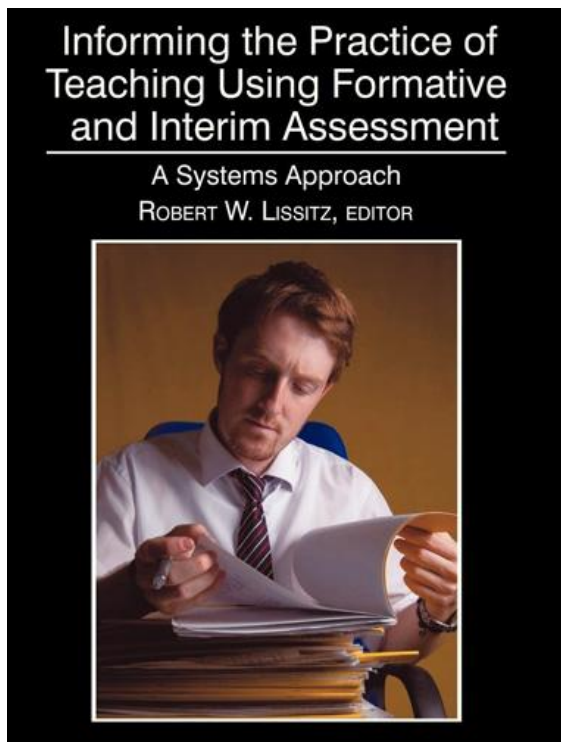
Before tackling the entire book, it might be constructive for those new to formative assessment to read two works cited in multiple chapters – the seminal 1998 "Inside the Black Box" by Black and Wiliam, and

the extensive 2001 National Research Council Report *Knowing What Students Know*, by Pellegrino, Chudowsky, & Glaser. Both are easily found via a web search. If one identified the most frequently cited works within the 11 chapters of the book, that itself might provide a reading list for a course called Best Practices in Classroom Assessment.

I begin with Chapter 2, entitled "Design of interim assessment for instructional purpose: A case study using evidence centered design in Advanced Placement," authored by Lori Nebelsick-Gullett, Cindy Hamen Farrar, Kristen Huff, and Sheryl Packman, because it provides clear descriptions about how the College Board used formative assessment to roll out the latest revision of the Advanced Placement Biology exam. The principles of Evidence Centered Design, which are well described, and AP Insight, which "...is a system designed to provide both high-need students and teachers the support they need to overcome the current challenges in the AP classroom (p. 28)" combine multiple elements that implement an array of activities illustrative of years of research on formative and interim assessment best practices.

As the College Board developed the new AP Biology exam, they focused on the big ideas of biology, wrote learning objectives for essential knowledge statements, designed science practices that would assess student understanding of the core ideas, and developed assessments that would *inform the teacher and the student* about the next instructional steps required for mastery. A key to this process is the supportive materials for the teacher: "Throughout the program, teachers are provided assessment strategies, instructional resources, *and professional learning experiences to support their ability to mobilize the formative*

process in their classrooms: and to understand and recognize evidence of student learning (p. 31, emphasis mine.)." With considerable background information on actionable feedback and its role for the student and the teacher, the AP Insight system builds feedback capability into the interim assessments within the AP Biology syllabus to provide teachers with clear directions on how to move students over the conceptual misunderstands that are discovered in the assessment process. This dedication to building the pedagogical content knowledge of teachers as a design principle in the AP Insight system is how College Board takes formative assessment from the theoretical into the practical – it is a wonderful illustration of making formative and interim assessment a real driving force for student and teacher success. It's also the best chapter to inform the practice of teaching.

Of course, unlike most school districts or high school science departments, the College Board has the resources and the talent to build such a complete instructional and assessment system. The principles of an effective instructionally oriented balanced assessment system are all here, and go well beyond the curriculum development, assessment training, and feedback loops commonly built into most formative assessment programs. The explanation of supports for improving the teacher's ability to use the assessment results to retrain and refocus his/her own next steps sets this chapter's discussion apart from the rest of this book.

In Chapter 1, "Connecting the dots: Formative, interim and summative assessment," Dylan Wiliam, Gage Kingsbury and Steven Wise offer a substantial introduction to the principles and research behind the concepts covered in the

rest of the book. They suggest that the current emphasis on summative assessment provides an overall look at the achievement of large groups of students, but is not useful to teachers or students because information about student progress comes too late, provides little actionable information about individual student needs, and offers no information to parents about how to support student learning. Their explanations of summative, interim, formative, and classroom assessment, with illustrative examples of how questions can be classified by their intended use, and their discussion of test validity, offer a sound introduction to the principles addressed by the remainder of the book.

Chapter 3, Integrating student standardized and formative assessments with the National Board for Professional Teaching Standards' Teacher Development Process, by Joan Chikos Auchter, offers an interesting history of the traditional National Board Certification process and a new program (2006) called NBPTS *Take One!* that offers schools the opportunity for teachers to do a school improvement project based on a lite version of NBPTS standards. Auchter works hard to illuminate how NBPTS' core propositions for teachers often require the use of formative assessment to enable teachers to 'recognize individual differences in their students and adjust their practice accordingly' (Core 1), to 'regularly assess student progress' (Core 3), and to incorporate education research and participate in learning communities (my connection). These elements reflect best practices in using assessments to improve teachers' own abilities to address student learning needs.

Her report on the first year of a three-year pilot study includes an element of formative assessment that she does not identify as

clearly as she might. In the NBPTS process, teachers are evaluated by expert teachers on multiple facets of their professional performance – on all five of the core principals. They receive clear feedback from these observers designed to guild their own improvements in the classroom. This goes well beyond student achievement measures, and gets into the multiple measures for classroom improvement that the research suggests is critical to school improvement. And this is where, I think, the NBPTS *Take One!* program really does incorporate formative principles. It's the feedback to the teacher that is formative, and is likely more important to the teacher's professional growth and the subsequent improvement of or demonstration of best practices that NBPTS is all about. While Auchter doesn't make this point, the reader comfortable with formative assessment and multiple measures can fill in the blanks and recognize the merits of the NBPTS efforts.

Chapter 4, "Using assessment data in real time: What teachers need to know and be able to do," is by another name familiar to K-12 assessment leaders, Margaret Heritage. Her work has trickled to the school level over the years, particularly as it relates to learning progressions for instruction. This chapter gets directly into what teachers can do with post hoc and real time assessment data to guide next steps in the learning process, and the concepts match up well with the many definitions of formative and interim assessment outlines by others throughout the book. She builds her chapter around principles of real-time assessment as advanced by the National Research Council 2001 publication *Knowing What Students Know.*

Her premises: An evaluation system should be comprehensive, coherent, and continuous. Comprehensive implies

"multiple sources of evidence to draw inferences about an individual student's learning....and inform ... the decision that teachers..." need to take (p. 87). Coherence means the system is based "on a shared model of learning that links curriculum, instruction and assessment within the classroom and links classroom assessment with external, large-scale assessments. Models of student progression in learning, the pathways by which students develop increasing expertise in a domain, should underlie the assessment system (p. 87)." Continuity means continuous assessment of student progress over time so that "student learning can be observed and interpreted (p. 87)." This is as good a conceptual framework for formative assessment in practice as one might desire.

She notes that appropriate use of assessments supports teacher knowledge and skills – and goes back to the *Knowing What Students Know* by noting there were clear recommendations there to develop "user-friendly descriptions of how students learn content, which identify important targets for instruction and assessment (p. 99)." This kind of work would describe student behaviors during the learning process and clarify for teachers the sequential instructional steps necessary for student success. This, in turn, provides a basis for the teachers' own learning: "American teachers need to have the opportunity for deep learning and reflection with their peers...(p. 101)" as teachers in more successful national systems enjoy. Further, she notes that providing teachers with opportunities to understand the effective use of real-time data is sorely lacking, "underdeveloped and generally undervalued in the United States (p. 101)."

Readers who desire solid background information that is immediately useful for the classroom would do well to seek out more of Heritage's scholarship.

Chapter 5, "The instructional influence of interim assessments: Voices from the field," by Lisa M. Abrams and James H. McMillan, looks at interim assessment "located between teachers' minute-by-minute and daily formative assessment...and the summative unit assessments or test conducted after instruction... (p.106)" and its formative applications. Noting the blurring of definitions of formative and interim assessments arising from the proliferation of testing companies who are offering products designed for use throughout the school year, the authors draw on the CCSSO 2008 definition and note the language used by practitioners that formative assessment is 'assessment *for* learning' that emphasizes the use of data to inform instruction. Where those assessments are farther removed from the daily monitoring of student progress but still designed as data to improve instruction, the concept of interim assessments used formatively enters the data driven instructional world.

With NCLB and the increased use of assessments as measures of teacher and school performance, the authors note the significant increase in the use of interim assessments across the nation. The literature on how such assessments are being used is sparse, but the few studies cited reference legitimate formative strategies, such as identifying individual student weaknesses on standards, identifying error response patterns that signal instructional gaps, grouping students for additional instruction, identifying curricular gaps, and even identifying professional development needs for teachers. However, other researchers are cited who suggest teachers are not making

very good use of the data – interim assessments are not changing the way teachers reteach, nor are teachers engaged in "deeper conversations about instructional content and learning processes (p. 112)."

Reporting on two of their own studies, they confirm many of the limitations cited when using interim assessments to improve student performance. In the first study, teacher interviews identified several themes suggesting formative applications of interim testing, but the authors conclude "...there is only moderate evidence to suggest that they using [sic] results in ways that would be described as formative. While there was clear potential for using results in formative ways, the most significant constraint for doing so seemed to be the lack of time – time to thoughtfully analyze data, time to meet and collaborate with colleagues, and time to provide high-quality instructional correctives (p. 120)."

Their second study used a 176-item survey and analyzed relationships between factors and contextual conditions influencing the use of interim test results. They found most of the elementary teachers used interim assessments to alter instruction, but the depth and the nature of the adjustments were less clear. The best predictor of the use of data for instructional change was the time spent on data analysis, and an inhibitor was lack of time to do this kind of work.

They conclude that it appears clear teachers are using interim assessment data as a driver of some level of instructional modification, though it generally does not appear there is substantial depth in how that is being done. Their recommendations make sense, and include: provide teachers with test questions and answer options; allocate time during the school day for collaborative analysis; organize professional learning communities

to discuss assessment data; address the tension created by curriculum pacing pressures that reduce the opportunity to reteach or remediate; and enhance teacher expertise in data analysis and interpretation. These suggestions are extremely challenging, if not impossible, where interim assessments are copy written and secured publisher tests.

Chapter 6, "Sourcing instructionally embedded formative assessments," by William D. Schafer, offers a utopian but appropriate and logical vision for the development of formative assessments that are fully integrated in instructional units. Schafer offers an action plan for creating instructionally embedded formative assessments. Multiple steps would be involved. Groups of teachers trained in formative assessment and curriculum would develop curriculum units that include multiple assessment formats: expert teacher review and tweak the units and assessment designs: in-service instruction be offered to teachers to prepare them for the effective use of formative assessment techniques: the units be piloted and become available in searchable databases. States should improve the elaboration of their summative assessment design – outlining 'the content, process, and difficulty distributions of the items on them (p. 141)." The blueprint would identify the range of topics and activities expected from students and clarify what students are expected to demonstrate. Teachers need to be trained to use assessments to motivate students and help students make judgments about their own instructional needs. Finally, teachers need to see many examples of "assessments designed for formative purposes."

This outline is just the kind of work suggested for teachers by the body of teacher training materials available from

some of the other chapter authors – Brookhart, Stiggins, and Heritage come to mind here. And I suspect both the two national assessment consortia believe they are engaged in just this kind of work, though their connection to teachers is so far rather remote. However, it appears the vision Schafer suggests is slowly emerging in many venues, and one can hope to see it continually evolve.

Chapter 7, "Marrying formative, periodic, and summative assessments: I do," by Kimberly O'Malley, Emily Lai, Katie McClarty, and Denny Way, offers a variation on the term interim assessment. Periodic assessments, they suggest, "differ from formative assessment in two key ways. First, their purpose goes beyond informing instruction. They are used to evaluate student progress on learning the content standards, to assess student mastery of a specific content domain, and to predict performance on the summative assessment. Second, they occur outside instruction (p. 147)." They then divide periodic assessment into two types: Domain specific, such as a test of statistic and probability in grade 8 math curriculum, or cross domain, which assess performance on all domains of a content area to provide information on the full content area.

They suggest that optimal use of formative, periodic and summative assessments in combination offers a "comprehensive system that can be used to support the teaching and learning process throughout the year (p. 148)." They offer three considerations for the design of such a system: timing; the kinds of information provided by the three assessment types, and reporting and feedback options.

They outline the types of assessment data that would be most useful for schools as the

year progresses from fall, into winter and spring, and their discussion of what to do with the accumulated data in the summer offers a much-needed missing step in the usual application of assessment data to drive instruction.

Their discussion of the types of information needed is also noteworthy, particularly because they explicitly identify both teacher data needs and student data needs. This is a critical element in data-driven instruction and, as they point out, in the theory of action promoted by Black and William, a necessary driver for successful use of formative assessment to involve students in their own instructional planning.

Also pleasing is the discussion of the importance of reporting and feedback, their third consideration. They stress the importance of feedback that is "timely, clear, and actionable (p. 159)" and that is directed to both teachers and students, again stressing the role of active student participation in the process. Notably, they identify a significant shortcoming of summative assessment – the results come too late and are often too generalized to be useful to either students or teachers. They describe, with some optimism, the arrival of online testing that offers the opportunity for summative assessment feedback to be rapid and offer enough detail to actually be immediately useful in planning next steps.

They also note the *utility of having access to summative test items*. There seems to be little chance that the two national testing consortia will make many items available annually, but doing so would be welcomed by those of us who have been doing this kind of assessment professional development in the field.

School districts and teachers who are using multiple interim/periodic assessments, or considering doing so, should study this chapter. Few districts implement the comprehensive outline here: Where interim assessment are in place, there's usually very little effective formative assessment to flesh out instruction, and most of the authors of this book would suggest formative assessment is critical to any strong assessment program. Districts might check their own assessment schedule and implementation plans and consider adding the elements suggested here if they are missing.

Susan M. Brookhart writes "Comprehensive assessment systems in service of learning: Getting the balance right," in Chapter 8. She is the author of several books on assessment aimed at K-12 teachers, and well worth adding to one's assessment library. Her review of balanced assessment literature and her suggestion for a different model begins with a good review of the typical tripartite model of balanced assessment systems – formative, interim and summative assessments, and one in which there are multiple uses for interim assessment. She identifies two "stress fractures" in this model – the lack of clarity over the use of interim assessments – are they instructional, evaluative, or predictive, and the lack of clarity on where classroom unit tests fit – are they summative, interim, and/or formative?

She notes that the two national testing consortia are both expressing consistent descriptions of interim assessments as instructionally supportive and not evaluative, which she sees as positive in the evolution of assessment systems. However, she also notes the lack of clarity about where classroom unit tests (generally summative assessments) and grades fit in a balanced assessment model, despite their frequency of

occurrence and therefore their importance in any evaluation model. Her proposed new balanced assessment model is arranged in four quadrants instead of the usual tripartite assessment triangle. On the x-axis one finds assessment purpose – to the left is more formative purpose and to the right is more summative purpose. The y-axis is the test scale – up for large scale and down for classrooms. So the upper left quadrant has interim/benchmark assessments, lower left is classroom formative assessment, upper right is large scale accountability tests whether at the state or district/school level, and lower right is grading and classroom summative assessments.

With this focus on purpose of assessments and the scope of the assessment, she provides a useful visualization of a balanced assessment system that includes the missing grading dimension that remains so troubling in any teacher professional development session on formative assessment. Ultimately, teachers have to provide some kind of grade that goes on a student record, and conceptualizing grades as a part of balanced assessment should be helpful as the debates over grading become more vocal.

Chapter 9, "Errors in student learning and assessment: The Learning Errors and Formative Feedback (LEAFF) Model," by Jacqueline P. Leighton, Man-Wai Chu, and Paolina Seitz offers a useful perspective on learning errors and their role in the learning process. They begin with a brief overview of the literature on errors and note that errors are an integral part of learning. Differentiating between novice errors (resulting of lack of knowledge) and expert errors (lapses in memory due to fatigue, not lack of knowledge), they suggest "Errors facilitate learning by acting as opportunities for a mentor or instructor to provide formative feedback – that is feedback that is

nonevaluative, supportive, timely, and specific.... and perceived as relevant by the learner (p.187, from Shute, 2008)."

They cite other research on the mastery-oriented learner vs. the performance-oriented learner, in which the former seeks mastery learning and is motivated to try new means to reach success, while the latter seeks praise and a good grade and is less inclined to seek new learning situations in which success is more risky. Learning for the performance group is a means to and end – praise, and not new knowledge and skills. So the challenge is finding ways in which all learners can view errors as opportunities to grow through effective formative feedback.

The LEAFF Model is proposed as a process based on the research they have cited which does three things to improve student performance. First, it focuses on the instructional climate, when teachers create learning environments "that either promote safety for learners making mistakes or promote danger for learners making mistakes (p. 199)." This involves explicitly discussing the value of errors, and understanding that errors are opportunities for learning. Second, and the basis of the LEAFF model, is supporting student *mental models* of learning – helping students build positive perceptions about the learning environment that will make them receptive to feedback and learning at a deeper level. The model offers two paths for teachers in order to accommodate students who see errors as opportunities to move forward and students who "negatively appraise the learning environment as unsafe and, therefore avoid making mistakes in the service of showing what they are learning (p. 202)." Finally, the model focusing on student performance, and suggests that performance will be higher among students who see the learning environment as safe

compared to those who see the environment as unsafe.

This construct is consistent with those who support constructive feedback designed to give the student clear information about next steps to learning, and this is a fundamental construct underlying formative assessment. The emphasis in LEAFF, however, is grounded in helping all students understand the potential of seeing error analysis as an opportunity for learning, and moving the classroom climate from one in which errors are risky to one in which errors are an important step in making progress toward mastery. The authors suggest further research on their model. This is a good suggestion – the model might serve as a framework around which schools could build implementation plans for any well-designed formative assessment process.

Chapter 10, "Defining systems for learning," by Paul Nichols and Charles DePascale, describes the design of a computer "System for Learning" simulation designed to model a four-part system to improve student learning. They begin by offering background on educational systems as described by other scholars, each of which emphasized different elements of their individual systems. They then review briefly the elements of systems thinking and system development, and proceed to outline and explain the elements of their own model. They suggest a three-dimensional diamond as their framework – with a Theory of Learning at the apex, and the other 4 points of the diamond being curriculum, assessment, instruction, and professional development. They further break down assessment into classroom, interim, and large scale assessments. They don't offer details or specifics of how to operationalize their four elements – they don't suggest a specific Theory of Learning, for example,

leaving that to those who might want to use their generic model to measure/model outcomes.

Following the background discussion, they describe the technical elements of their model – a quantitative model using dynamic modeling software, specifically a visually oriented graphical simulation program called STELLA. STELLA has three kinds of variables: state, flow and converters. For those who have a background in statistical modeling, the outlines of how these elements are defined and the roles they play in the model will be clear. Running the simulation for various time periods produces information about the effects of manipulating various elements, for example professional development, exploring innovations, implementing innovations, implement the Associationist Approach Theory of Learning (their example), and seeing their effects on numbers of students reaching proficiency over time.

This kind of modeling is, to my knowledge, rare in education. It is certainly not within the capability of most local school districts. Noting the challenges, the authors admit "The difficulty in furthering interest in and applications of system modeling is that understanding system behavior is prevented by the very specialties, such as curriculum specialists and psychometricians, on which the field of education depends. The information about the nature of system in education, how such systems develop and how such systems operate, is dispersed across these specialties (p. 223)." Systems thinking is not usually a part of undergrad teacher education or the education coursework required within most educational specialties, nor is quantitative modeling. States or perhaps large city school districts may have specialists capable of such work, and external evaluation firms

who monitor large program implementations may be able to model as well. How such modeling can inform education reform remains in the realm of theory.

Chapter 11 is entitled " Productive formative assessment *always* requires local district preparation", written by Rick Stiggins and Steve Chappuis, who are representative of the staff of the Assessment Training Institute, founded by Stiggins in 1992 to improve the use of assessment for learning. This chapter effectively summarizes actionable steps the authors believe are necessary to implement a formative assessment program within classrooms, and is an appropriate and ideal final chapter for the book.

The Assessment Training Institute has been offering very thorough and well-developed classroom and administrative training materials about formative assessment since its founding. Their best-know book, *Classroom assessment FOR student learning: Doing it right – using it well"* by Stiggins, Steve and Jan Chappuis, and Joan Arter, is still the most thorough 'how-to' manual for teachers, and provides the theoretical background on formative assessment, an in-depth look at multiple assessment formats, and detailed examples of how to design classroom formative assessments to generate actionable feedback that involves students in identifying their own next steps in learning. If one seeks to find a single resource for bringing formative assessment alive in a K-12 classroom, this seminal text is where to begin.

They argue "Essentially, all assessment is local. Regardless of the level of test administration – no matter where the scores come from – they feed into the local instructional decisions that determine school effectiveness (p. 238)." They note that the

two assessment consortia are offering "at least a nod in the direction of formative applications of assessment (p. 238)." And they argue, as the Assessment Training Institute and Stiggins have argued for 20 years, "The time has come to abandon the belief that annual standardized test scores provide sufficient information to support the development of effective schools. They have not for 60 years and will not in the future because they fail to meet the information needs of crucial instructional decision makers (p. 239)."

One with some familiarity with formative and interim assessment concepts can match content from the other chapters of this book with each of the action elements Stiggins and Chappuis outline here. Action 1 is building balanced assessment systems, and their chart of the ingredients of such a system of classroom, interim, and annual testing offers clear descriptions of key issues and formative and summative applications of all three assessment types.

Action 2 suggests continuing to refine assessment expectations, characterized by the recognition that standards alone are not sufficient to support school improvement. Standards should be "arrayed in learning progressions, so they unfold .... in a manner consistent with the way learning happens (p. 241)." And "Each standard must be deconstructed into the scaffolding student climb to get there... and transformed into student- and family-friendly versions to facilitate understanding and learning (p. 241)." This is the concept of learning targets that Stiggins and others promote in building systems that parse learning into segments that can be clearly measured and permit clear feedback to teachers, students and parents.

Action 3 is ensuring assessment quality, and the authors note that most local assessments

are weak and teachers simply are not trained in effective assessment design. Good assessments require clear purpose, clear learning targets, an understanding of how to devise an assessment method to match the learning target, and an understanding of how to communicate results and next steps.

Action 4 suggests learners need to become assessors. Bringing students closer, and into the assessment process is a key to helping students make appropriate decisions about next steps in their learning. Otherwise, assessment is something we do *to* students, not *for* students, and assessment *for* learning is their definition of formative assessment.

Action 5 is to drop our belief that grades and annual test scores communicate anything that supports student learning. They are "important periodic judgments of student learning (p. 244)" but don't help the learner know what to do to improve. Feedback to support learning is descriptive, not judgmental.

Action 6 suggests that learning success motivates students. Intimidation derived from accountability is not motivational. "Upon seeing the results of their own classroom assessments, we must bring student to a place where they can say: 'I understand these results and I know what comes next in my learning (p. 245).' "

Action 7 is increasing assessment literacy – for teachers and school leaders. "Very few have even been given the opportunity to become assessment literate p. 245)" and are therefore unprepared to effectively create or implement a balanced assessment system.

Were I designing a pre-service course in effective classroom assessment, the collective work of Stiggins and the Assessment Training Institute, now a part of

Pearson Education, would be my primary resource and Stiggins and the ATI staff my mentors. They have taken years of assessment research and developed operational manuals for putting effective assessment practices into practical classroom formats, and they offer substantial specific information targeting administrators who want to create formative cultures within their schools. I would add other names here: Susan Brookhart from this collection with many easy to use formative strategies for teachers, Margaret Heritage and Karin Hess on learning progressions. Jay McTighe, Robert Marzano, Debra Pickering, Nancy Frey and Douglas Fisher, and Edie Holcomb come to mind as authors with practical books on improving student achievement through effective use of assessments and student performance data. Each of these authors provides concrete classroom-based means to implement effective assessment techniques in classrooms.

Two shortcomings of this edition were unfortunate. First, the reproductions of graphic images were generally fuzzy, and in several cases distractingly so. Charts and tables produced by word processors or spreadsheets were fine. And a few chapters would have benefitted from a simple spelling and grammar checker that would have identified double words, misspellings, and syntax errors that should not have been so prominent.

This collection reflects the work of supporters of formative and interim assessments of student learning as a school improvement process. I am also a strong proponent of formative and interim assessments as tools to provide actionable data to teacher and students that immediately informs next steps for the teacher and learner. It simply makes instructional sense in every manner in which

I have experienced it, and taught the principles to others.

 On the other hand, a four year study entitled "Classroom Assessment for Student Learning: Impact on Elementary School Mathematics in the Central Region: Final Report," published in 2011, is not mentioned by any of the authors, but is an in-depth examination of formative assessment in practice. Coming from the Regional Education Lab, Central, it is online at http://ies.ed.gov/ncee/edlabs/regions/central/pdf/REL_20114005.pdf. Because it focused on a significant multi-school program with extensive teacher training and ongoing support from the Assessment Training Institute, which I have already praised, I tracked its progress and contacted the principal investigator to let me know as soon as the study completed its peer review. The results were, in two words, surprisingly disappointing. Karee E. Dunn and Sean W. Mulvenon in a 2009 article from the online journal *Practical Assessment, Research & Evaluation* entitled "A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education," suggest there is but a limited body of empirical evidence supporting formative assessment, from Black and Wiliam in 1998 through the best research available at their publication date. Taken together, this scholarship offers some pushback to the increasingly strong support for formative assessments emerging today, and implies the necessity of moving carefully in this direction in light of the enthusiasm for new assessment models that pervade K-12 education today.

When we see teacher and administrative preparation programs at all levels regularly incorporating the concepts of balanced assessments as tools for educational success, efficacy research can be done on a large scale to improve our understanding of how well balanced assessment actually works. In the meantime, *Informing the practice of teaching using formative and interim assessment* is a useful look at the topic by promoters of balanced assessment systems.

About the Reviewer

Brian Preston retired from regular service in education in August, 2012, after 45 years in the field. He spent 20 years in middle and high school classrooms in Michigan, Germany, and New York, in both public and private schools. He completed his administrative doctorate at NYU and went into administration first as a 6-12 assistant principal and Social Studies Department Chair. He ran a UAW-GM Education Center within a General Motors Assembly Plant, and coordinated Institutional Research and wrote regional state and federal grants at the Nassau County BOCES in New York. There he wrote and then served as project director of a 5 year $15 million technology professional development grant for 128 districts on Long Island, NY. The grant included data analysis for instructional improvement, which began a major shift in the focus of the remainder of his career. He closed his career with the Lower Hudson Regional Information Center, where he supported the integration of data warehousing services and instructional support programs, and continued working with teachers and administrators interested in using assessment data formatively. He has taught in graduate education programs at New York University and the New York Institute of Technology, served on the Board of the Northeast Educational Research Association, and was one of the 12 founders in 2000 of the New York Schools Data Analysis Technical Assistance Group.   Currently, he comments on national educational policy issues at www.k12edtalk.com.

Follow *Education Review* on   Facebook and on Twitter: #EducReview