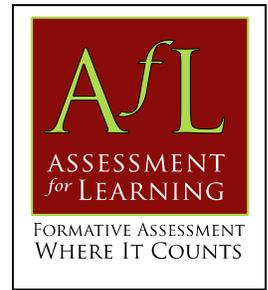


DEFINING AND ENHANCING FORMATIVE ASSESSMENT

W. JAMES POPHAM
UNIVERSITY OF CALIFORNIA, LOS ANGELES
SEPTEMBER 15, 2006



Preamble

The following analysis is separated into two parts. The first section represents an attempt to provide those individuals attending an October 10-13, 2006 meeting in Austin, Texas with a proposed definition of formative assessment that, either as presented here or in a subsequently modified form, can be accepted by those present. The Austin session is the initial meeting of a State Collaborative on Assessment and Student Standards (SCASS) sponsored by the Council of Chief State School Officers (CCSSO). The focus of this SCASS is to be Formative Assessment for Students and Teachers (FAST). In my view, this new FAST SCASS activity represents a potentially important initiative which, if effectively implemented, can have both a short-term and long-term positive impact on American schooling. This new SCASS is intended to enhance the caliber of district-level and classroom-level instruction in a manner rarely undertaken in the past, at least collaboratively, by the nation's chief state school officers. The focus of the new initiative is to be "formative assessment." It is apparent, therefore, that all FAST SCASS participants' understanding of what is actually meant by formative assessment should be identical.

The initial section of the following analysis, then, provides a proposed definition of formative assessment for the Austin group's consideration. I take no credit for the proposed definition. On the contrary, over the past three months, largely in preparation for the upcoming Austin FAST SCASS meeting, a large number of individuals, myself included, have been exchanging electronic views regarding how best to define formative assessment. The hard-wrought consensus of those e-mail deliberations will be presented in Section One of this paper.. I am most likely overlooking someone who took part in fashioning this proposed definition, but among those individuals who had an electronic hand in shaping the definition of formative assessment to be proffered here were Dylan William, Lorrie Shepard, Rick Stiggins, Scott Marion, Phoebe Winter, Don Long, Stuart Kahl, and Brian Gong. Please be assured that the definition to be presented below reflects not only my view but, rather, the composite counsel of the foregoing individuals as well as others. I will conclude Section One of the paper with an attempt to explicate more fully the key components of our group-proposed definition. It will be the decision of the Austin FAST SCASS participants about how to finally formulate the definition for formative assessment that will guide the efforts of this new CCSSO initiative. I hasten to add that the definition of formative assessment to be offered here represents, in my opinion, a useful way for us to characterize the focus of our FAST SCASS endeavors.

Section Two of the paper represents my own recommendations regarding the way to enhance the effectiveness of any formative assessment. It is one thing to create a formative assessment. It is quite another to create a formative assessment that actually contributes to improved instruction and, as a consequence, promotes better student learning. I will conclude Section Two by laying out several suggestions, born of my own experience—sometimes successful, often the opposite—regarding how to render formative assessments optimally useful to both teachers and students. I would love to be able to blame any shortcomings of the paper's second section on the individuals who contributed to Section One's proposed definition of formative assessment. But, regrettably, I can't. Any deficits in Section Two are mine alone.

Section One: Defining Formative Assessment

The words “formative” and “summative” have been in English-language dictionaries for well over 100 years. My current dictionary defines *formative*, first used in the early 19th century, as an adjective indicating that whatever noun it modifies is “capable of alteration by growth and development.” *Summative*, whose first recorded use in English apparently occurred about 1880, is defined as an adjective whose modified noun must be “additive” or “cumulative.” Yet, despite their long lineage, these two terms never made all that much difference to educators until 1967 when Michael Scriven drew his classic distinction between “formative evaluation” and “summative evaluation.” Although Scriven’s (1967) formulation of two different uses of educational evaluation was stimulated by the evaluative requirements of a U.S. federal law, the Elementary and Secondary Education Act of 1965 (ESEA), and thus at the outset was of most interest to U.S. educators, Scriven’s distinction between formative and summative evaluation has been accepted worldwide.

Because 1965’s ESEA provided for unprecedented distributions of federal dollars to state and local U.S. educators, the provisions of that law were obviously of considerable concern to American educators. One of the key requirements of this new law was that, in order to qualify for *next year’s* federal largesse, local educators who were recipients of *this year’s* ESEA funds were obliged to evaluate whether their current federally supported projects had been successful. Unfortunately, one reality of the mid-sixties was that American educators didn’t know squat about how to evaluate their educational programs as called for by ESEA. And so it was that a number of first-rate academics devoted their energy to promoting a better understanding of the nature of educational evaluation. This flurry of evaluation-related thinking stemmed, unarguably, from the need to satisfy ESEA’s evaluation requirements. Scriven’s formative/summative distinction was one of the earliest, and most influential, evaluation-specific constructs to be accepted by America’s emerging collection of educational evaluators.

Scriven envisioned two roles for evaluation. He defined formative evaluation as the appraisal of an educational program’s worth or merit while there was still time for the program’s staff to make modifications to improve their program. Summative evaluation, in contrast, was seen by Scriven as an appraisal of the worth or merit of a mature,

essentially final-version educational program. Thus, formative evaluation was intended to provide evidence to a program’s staff so they could make improvement-focused decisions regarding their still malleable program. But summative evaluation was to provide relevant decision-makers with the information they would need to make a “go/no-go” decision, that is, a decision to continue or terminate the educational program being evaluated.

Background

For fully 40 years, Scriven’s distinction between the formative and summative evaluation of educational programs has been a valuable way for many evaluators, not only in the field of education, but also in a variety of other arenas, to conceptualize the nature of their evaluative efforts. But educational *evaluation*, of course, is not educational *assessment*. Though these two terms are surely relevant to one another, they are importantly different. Yet, in recent years, “formative” and “summative” are being increasingly used as adjectives to modify not only the noun *evaluation*, but also the noun *assessment*. Accordingly, it has become apparent that educators need to be sure they not only understand the meaning of *formative assessment* and *summative assessment*, but also that they employ those two labels accurately.

Although, during the four decades between Scriven’s 1967 formative/summative distinction and the present, there have been occasional essays addressing formative versus summative assessment (e.g., Sadler, 1989), interest in formative evaluation among U.S. educators was spurred chiefly by two events.

First, in 1998 Paul Black and Dylan Wiliam authored an article in the *Phi Delta Kappan* that attracted considerable attention from U.S. educators. In that essay, these two British researchers reported results of a meta-analysis focused on the instructional payoffs of classroom formative assessments. Black and Wiliam concluded that the appropriate use of classroom assessments not only improved students’ learning of what was being taught in class, but such assessments also *bettered students’ scores on external achievement tests as well*. The well-crafted argument fashioned by Black and Wiliam soon began to reach the ears of educators, not only in the U.S. but in other nations as well. However, interest in the Black and Wiliam advocacy of formative

assessment was most certainly heightened in the United States by enactment of the No Child Left Behind Act (NCLB).

NCLB was signed into law by President George W. Bush on January 8, 2002. This potent law, coincidentally, was merely the most recent re-authorization of ESEA of 1965, the statute that had triggered Scriven's generation of the distinction between formative and summative evaluation. NCLB, however, called for a far more stringent form of educational accountability than had been required in either its 1965 version or by subsequent re-authorizations of ESEA. If public schools receiving NCLB dollars failed to improve their students' scores satisfactorily on state-selected accountability tests (required by NCLB at more than double the grade levels demanded in earlier ESEA incarnations), those schools could be designated as requiring improvement, re-structuring, or flat-out obliteration. Clearly, the pressure on U.S. educators to boost their student's scores on NCLB tests was profound.

In the midst of frenzied activities to elevate students' test scores, starting in 2002 and continuing to this day, many American educators began to recall the message of the *Kappan* essay of a few years earlier. It was an essay which said, in simple words, the incorporation of formative assessments was a research-supported way to improve students' scores on external achievement tests. And, of course, NCLB accountability tests were most assuredly the "biggest and baddest" external accountability tests American educators had ever seen.

Accordingly, like flies questing for honey, in-the-trenches U.S. educators began to clamor for "research-proven" formative tests that could help them dodge the accountability arrows being aimed at them annually by the need for schools to make "adequate yearly progress" in the form of improved NCLB test scores. Because America's test-development firms are not staffed by ninnies, some of those vendors began to market tests which they cunningly labeled as "formative assessments." In reality, the test vendors were altogether free to do so, because no reputable or quasi-reputable group had ever spelled out the defining characteristics of what truly made an assessment formative. Arguing by analogy from Scriven's evaluation distinction, it was widely (if loosely) thought that formative assessments had something to do with improving yet-malleable instructional programs, whereas summative assessments were to be used in making final judgments about a program's or a student's achievement. Examples of summative assessments

were the many NCLB-spawned annual accountability tests intended to determine how effectively schools, districts, and states were performing. But such a gross, very general distinction between formative and summative assessment was about all that most educators carried around in their skulls. Although Black and William had actually set forth a rather constraining definition of formative assessment in their 1998 essay, most American educators were operating on a far more relaxed notion about how formative and summative assessment actually differed.

However, because numerous test companies have recently begun to inaccurately cite the research reviews of Black, William, and others to indicate their newly labeled (or hastily created) "formative assessments" are demonstrably effective in raising scores on NCLB tests, CCSSO has launched a major initiative related to formative assessment. This initiative is intended to help the nation's educators better understand what formative assessment is (and is not), the nature and limits of formative assessment's empirical support, and how best to employ formative assessment to benefit students' learning. A key component of this major CCSSO initiative is to be the FAST SCASS. Obviously, for an initiative of this magnitude to have any meaningful chance of being successful, there must be agreement among those involved regarding what is actually meant by the phrase "formative assessment."

Definition Time

Frank Philip and Don Long of CCSSO, therefore, have encouraged several of us who had been advocates of formative assessment to engage in a continuing conversation, typically by e-mail but occasionally via conference telephone calls, aimed at reaching a serviceable definition of formative assessment. During the summer of 2006 those interchanges took place. There was substantial give and take regarding the pivotal attributes of formative assessment, particularly so that the resultant definition would be consonant with the findings of several meta-analyses supporting the instructional dividends of formative assessment. Numerous definitional modifications were proposed, exchanged, approved, rejected, or massaged. Presented below, then, is the definition of formative assessment we wish to present to FAST SCASS members for their consideration:

An assessment is formative to the extent that information from the assessment is used, during the

instructional segment in which the assessment occurred, to adjust instruction with the intent of better meeting the needs of the students assessed.

As you can see, there are several key components of this definition that should be isolated and, in some instances, clarified. I will try to do so now.

First off, when this definition uses the term “assessment,” it should be immediately recognized that this need not be a formal paper-and-pencil test or, increasingly these days, a computer-presented test. Rather, the above definition envisages the use of both *formal and informal* ways of gauging students’ status with regard to what is being taught. For instance, students might be asked to indicate the degree to which they understand a presented idea by using some sort of traffic-signal scheme (such as one involving colored paper cups) whereby a student-presented green symbol means, “I understand well enough to explain it to others,” a yellow symbol means, “I think I understand it, but I’m not sure,” and a red symbol means, “I definitely don’t understand it.”

As another illustration of informal formative assessment, frequent questions can be posed to students so that a teacher garners a useful idea about the degree to which an explained topic has been properly understood. Short tests of a more conventional sort, for instance, a five-item mini-test dealing with a key subskill that students must master en route to their mastery of a more terminal curricular aim, could also be regarded as formative assessments according to the definition given above.

Note also in the definition that the “information” yielded by a formative assessment, given the potential informality of the procedures used, need not be reported as the “number correct” or “percentage correct” scores we traditionally see accompanying formal tests. The information, as the definition says, is to be used to adjust instruction. If the nature of the information is sufficient to provide guidance in the adjustment of instruction, it need not be provided in the form of what we have come to expect as typical test-score reports.

Another key feature of the definition is its requirement the information obtained from the formal or informal formative assessment procedures must be used *during the instructional segment in which the assessment occurred*. In other words, results of the assessment need to be available to teachers (and, ideally, to students as well) while there is still

time to make beneficial adjustments to instruction and still time to carry out those adjustments. So, to illustrate, if a teacher were carrying out a three-week instructional unit, and a classroom test were administered at or near the end of the three weeks, this belated assessment could not be characterized as formative. Simply put, if there’s no time to make the changes in instruction that may be indicated by the test’s results, the test should not be regarded as formative.

You will see that the proposed definition calls for the intended purpose of the assessment to make instructional adjustments so as to benefit *the students being assessed*. This stipulation is intended to make our conception of formative assessment consonant with the research results indicating its dividends must be used to make adjustments in instruction (as noted above, while there is still meaningful instructional time remaining) for the students who were assessed. This proviso in the definition definitively indicates that the increasingly popular type of “interim” assessment administered by schools or districts every few months during the school year would be regarded as formative *only* if information provided by those assessments was returned to teachers in time for the information to be used to make meaningful adjustments in the instruction provided to those students who took the interim assessments. If the results of an interim assessment aren’t provided in time for instructional adjustments to be made for the assessed students, then the interim assessment is not formative. This exclusion from the formative family does not suggest that interim assessments are not potentially helpful to educators. The proposed definition, however, clearly precludes certain interim assessments from being regarded as formative.

Another pivotal point in the definition is that instructional adjustments are made “with the intent of better meeting the needs of the students assessed.” This aspect of the definition arose from an August 2006 e-mail interchange among Dylan William, Fritz Mosher, Paul Black, and Tony Bryk regarding whether a formative assessment did, in fact, need to actually alter instruction. In their 1998 *Kappan* essay, Black and William defined formative assessment as follows: “We use the general term *assessment* to refer 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. Such assessment becomes *formative assessment* when the evidence is actually used to adapt the teaching to meet student needs.” Note that in 1998

Black and Wiliam required that the information from a formative assessment actually be used to adjust instruction “to meet student needs.” In other words, not only were there adjustments to be made in the instruction, but those adjustments *needed to work*. In the definition proposed in this paper, the position is taken that adjustments in instruction *must occur*, and that the *intent* of those adjustments must be to better meets students’ needs (that is, must be intended to improve instruction and, thereby, enhance students’ learning), but the consequences of such adjustments need not, of necessity, be successful.

Fritz Mosher pointed out, in one of the August e-mails referred to above, that there is peril in defining an entity by the success of its consequences. Rather, he urged, we should define formative assessment according to its function, then employ research inquiry to determine the distinguishing qualities of effective formative assessment. Dylan Wiliam concurred with this position on the grounds that requiring a formative assessment not only to lead to the alteration of instruction, but also that those alterations must be successful constituted a “counsel of perfection.” Thus, in an attempt to make clear what the proposed definition calls for, it is that the information from a formative assessment does, in fact, lead to adjustments in instruction, and that those adjustments are clearly intended to better meet student learning needs, but the adjustments need not necessarily result in improved student learning.

Finally, with all the preoccupation regarding adjustments in instruction, it would be foolish not to concede that in some instances, probably very few, the evidence from a formative assessment might indicate that there were no adjustments needed *at all!* Clearly, few instructional plans are flawless, and fewer flawless instructional plans are implemented flawlessly. But it is certainly possible. Thus, if the information provided by a formative evaluation indicates “instruction is proceeding spectacularly,” then teachers may not need to make any instructional adjustments. More often than not, even in such rapturous instructional settings, at least one or two students will need to alter the way they are tackling a particular curricular aim, so some adjusting will be needed by those students. When thinking about formative assessment, it just makes sense to regard it as an enterprise focused on the adjustment of instruction because, experience sadly dictates, instruction almost always needs to be adjusted.

Well, that’s an elaboration on what seem to be the key components of the proposed definition of formative assessment. Hopefully, because of the magic of e-mail promulgation procedures, the definition will be considered both in advance of the October Austin meeting and, obviously, during the session itself. FAST SCASS members need to reach accord regarding the nature of what their SCASS is supposed to advocate.

Section Two: Enhancing Formative Assessment

In the remainder of this analysis, I want to identify four concerns that, if the developers and users of formative assessments keep in mind, can improve the beneficial impact on student learning of formative assessments. I realize all too well that one could go on and on about ways of sprucing up assessments of any kind. For example, many authors have written numerous books about the care and feeding of selected-response and constructed-response test items. Educational tests, whether formative or summative, obviously ought to be as good as they can be—simply bristling with reliability and capable of yielding only valid test-based inferences.

In this final section of the paper, however, I want to briefly deal with four issues I believe can make a

major difference in the educational payoff of formative assessments. In turn, I will address (1) the necessity of involving students actively in the pervasive use of formative classroom assessments, (2) the typical need to distinguish between formative assessments intended for teacher-use and those intended for student-use, (3) the need to construct formative assessments so the information they provide is maximally informative to the intended recipients, and (4) the advantage of having formative assessment’s locus of control situated as close to the classroom as possible. As I indicated in the preamble to this paper, Section Two’s observations are mine. Although I have been strongly influenced over the years by the assessment thinking of many of the individuals identified earlier, the four topics I will treat

in Section Two represent issues I regard as among the most critical considerations to be kept in mind when using formative assessments to improve student learning.

Student Participation in Formative Assessment

A number of individuals writing on the topic of classroom assessment have urged that teachers must make their students active participants in the learning process, not merely as receptacles but, instead, as self-directed partners. Indeed, Black and William (1998) make an eloquent plea for the meaningful involvement of students in classroom assessments, not in a competitive sense, but in the spirit of learning collaboratively and, insofar as possible, under each student's self-direction. But even self-directed learners need to know how they are doing, and that's where formative assessment comes careening onto the scene.

If I were to poll 100 experienced American teachers to find out how many of them really have even a remote idea regarding what formative assessment is, I'd be amazed if 50 of them could come up with a fog-free notion about the nature of formative assessment. But, because I am in a mood to be amazed, let's say that a full 50 of the 100 teachers described formative assessment in a manner more or less consistent with the definition proposed in Section One. I'd then be willing to wager lots of "previously owned" lesson plans that almost all of those knowledgeable 50 teachers would be thinking about formative assessment *exclusively* from a *teacher's* point of view. In other words, the 50 measurement-moxie teachers would believe the role of classroom formative assessment is to produce evidence for teachers so that those teachers could adjust their instruction and, thereby, teach kids better. But such an all-too-common idea about formative assessment really misses the mark. It misses the mark by about 50 percent. That's because roughly half the payoff from classroom formative assessment can, and should, come from involving students actively in the *instructional* use of formal and informal assessment. A continuing mission of such assessment should be to help students make more knowledgeable decisions regarding their current learning tactics. It's tough for a student to relax because, "I've mastered this skill," when a formative assessment supplies inescapable evidence that skill-mastery is simply not there.

If one significant role of formative assessment in the classroom is to help students learn more effectively and more efficiently, then teachers will need to

engage in serious re-thinking about the role most classroom tests play as students' grade-determiners. Although in U.S. classrooms there currently exists a time-honored tradition of using tests to compare students, it is fundamentally wrong-headed to try to use a test to help students guide their own learning while at the same time using the results of that test to grade or rank those students. Teachers, of course, must give grades. And tests—some tests—will most likely play a part in grade-determination. But the vast majority of classroom tests, the ones intended to help students better manage their own learning, should be non-graded. To differentiate the "learning-focused" tests from the "grading-focused" tests, all a teacher needs to do is inform students in advance by saying something such as, "Because of the need for me to supply grades, the upcoming classroom test on Wednesday, unlike our usual ones, will be used for grading purposes." It is imperative not to contaminate the learning dividends derivative from classroom formative assessments by the needless affixing of grades to students' performances on those assessments.

Making classroom assessments a pivotal part of the instructional process will not take place overnight. It will take teachers, and most definitely students, a good long while to accept the idea most classroom tests can function in a way intended to improve learning, not appraise students. Teachers will obviously need to let students know, in advance, which tests are intended for grade-determination. But such tests really should be very few in number. And if we are going to get the most instructional mileage out of our classroom tests, then we need to get cracking right away with this approach. Teachers need to tell their students that "There's a new testing game in our class, and its *only* function is to help you learn better the things you need to learn!" I suspect you can see that if non-graded formative assessments are made a constant component of what takes place in a classroom, and the thrust of all such formative assessment is to help kids learn, a decisively different atmosphere ought to prevail in that classroom. Students can, sans embarrassment, reveal they don't know something. That's because there is a clear commitment on the part of the teacher to help all students learn what they need to, and the heart of this learn-better game is formative assessment. There will be a new zeitgeist present in any classroom dominated by frequent formative testing. It is a zeitgeist to be cherished.

Leahy et al. (2005) recently set forth a series of excellent suggestions regarding how to get students actively involved in the instructional use of formative

assessments. But the overall point I am trying to make here is a more fundamental one than *how* to involve students beneficially in classroom assessment. This point, simply put, is that if we fail to make students active participants in classroom uses of formative assessment, we will lose many of the classroom payoffs for formative assessment. I am worried that in the rush toward teacher-usage of formative assessments, we may overlook the enormous learning dividends available if we can only cut students in on the assessment action.

Different Users Often Warrant Different Assessments

In a number of instances, the same formative assessment can supply the information needed by both teachers and students—but not always. There is a fundamentally different reason that teachers and students require the information supplied by a formative assessment. Teachers need the information to help them decide whether they need to modify their instructional activities and, if so, what the nature of such modifications should be. Students need the information to help them decide how they are doing in trying to achieve a given curricular aim (such as mastery of a cognitive skill or a body of knowledge). Students need to know if their learning tactics are working and, if they aren't, what sorts of changes in those learning tactics might prove beneficial. As I suggest, sometimes the information-needs of teachers and students are coterminous, but often they aren't.

Thus, when building a formative assessment, the developer of that assessment should always be asking, "Is the information to be supplied by this assessment intended for teachers, for students, or for both?" If we are to get the most instructional mileage out of formative assessments, then in many instances those assessments should do much more than simply answer the question of whether students seem to be "getting it." Ideally, teachers should be able to gain insights about *how to adjust* their instructional activities. Ideally, students would be able to gain insights about *how to alter* their learning tactics.

Properly constructed formative assessments will yield readily interpretable information. There's no need to make formative assessments so complicated that their results require a code-book to decipher. Thus, in many cases there will need to be two variations of the same formative assessment, possibly similar for the most part, but different in meaningful ways—one for the use of teachers and

one for the use of students. My recommendation regarding the construction of a formative assessment is that there always should be a specific *and exclusive* focus on the intended users (teachers and/or students) so that the assessment approach can be crafted to maximize the ease and usefulness of the information provided by the assessment to the intended users.

Maximally Useful Information

As suggested above, the information produced when students complete a formative assessment ought to be easy to interpret and unarguably useful. What this means, of course, is that the person constructing a formative assessment needs to be mindful of what sorts of "make better" instructional options exist. For both teachers and students, those improvement options can often be illuminated by the early-on identification of the subskills and enabling bodies of knowledge a student must acquire on the way to mastery of the curricular aim then being sought. That identification of "what's needed along the way" can be based on a task analysis (often referred to these days as a *learning progression* or a *progress map*) in which the key subskills and/or significant bodies of enabling knowledge are delineated. Developing a first-rate task analysis to isolate what en route learning underlies students' acquisition of a worthwhile curricular aim is far more difficult than often thought. Nonetheless, the identification of necessary subskills and enabling knowledge is often the only way that formative assessment can be really helpful to either teachers or students.

To illustrate, if a task analysis reveals there appear to be three pivotal subskills a student needs to have first mastered in order to achieve a more lofty curricular aim, then a formative assessment could be built containing a sufficient number of items focused on each of the three subskills. After students had taken the assessment, it would then be possible to identify if a student were having difficulty with one or more of the subskills. For example, if selected-response items were being used in a formative assessment, one obvious way of isolating the nature of students' difficulties would be to construct each wrong-answer option so it arises from a specific, instructionally addressable misunderstanding. To construct a formative assessment, of course, without deliberately incorporating sufficient per-subskill items would miss an opportunity to help identify *where* students are having problems.

In some cases, if there were several subskills and several relatively discrete bodies of enabling

knowledge, separate formative assessments might address each of these precursors in isolation. There is no must-follow rule about how to formulate a formative assessment with respect to the gaining of information about students' en route progress. The nature of curricular aims varies substantially, especially with respect to the grain-size, that is, breadth, of the skills and/or knowledge those aims represent. But there is a must-follow rule with respect to the utility of the information yielded by a formative assessment, namely, that *the information must have a clear decision implication for its recipient*. Therefore, when formative assessments are being devised, the deviser must always be thinking, "How can I create this assessment so the action implications are as evident as possible to the recipient of the results?"

Whereas, I suppose, it is possible to create useful formative assessments without undertaking a task analysis related to the chief curricular aim currently being promoted, this seems mighty unlikely to me. The person generating a formative assessment must constantly be wearing an "instruction-focused" thinking cap if the assessment is for teachers and a "learning-focused" thinking cap if the assessment is for students. Clear thinking in the development of formative assessments is almost always benefited by the use of rigorous task analyses.

Locus of Control Considerations

The closer that formative assessments are to the actual instructional events taking place in classrooms, the more likely will be their positive impact on student learning. Teachers who volitionally choose to make formative assessments an integral, ongoing component of their instructional activities will be more inclined to make warranted adjustments in their instruction based on the results of such assessments than they would have been if they had been forced by superordinates to formatively assess their students. People generally are inclined to do with more enthusiasm what they choose to do, not what they are required to do.

Accordingly, we need to incline teachers to volitionally infuse frequent formative assessments into their instruction. And this will often mean that a state or a school district should supply teachers with already-constructed formative assessments for their use—if those teachers choose to do so. Let me be candid. The construction of instructionally illuminating formative assessments is far from fools' play. Busy teachers, although they might be bright enough, and motivated enough, and even skilled enough, will

rarely have *time enough* to generate first-rate formative assessments. Remember, in many instances there will need to be differences between formative assessments for students and formative assessments for the teacher. Moreover, a really useful formative assessment should be crafted so that it provides a reasonably clear, task-analytically based answer to the question: "What should I do next?"

Thus, a potentially useful role of a state department of education might be to generate crackerjack formative assessments dealing with as many significant state curricular aims as possible, then make those assessments available to that state's educators for *optional* use. If the assessments are really good, and can inform teachers and/or students regarding suitable next steps, it is likely most teachers will want to use the assessments. And please don't forget that formative assessment can involve very informal procedures. So, a state department of education might, in addition to distributing actual "tests," could also prepare guidelines and examples of the sorts of informal formative assessments that, in connection with the promotion of particular curricular aims, might be helpful.

Short-cycle, quick-turnaround formative assessments, for both teachers and students, have been empirically shown to create a beneficial impact on students' learning. We should urge teachers to employ numerous instances of formative assessment in their classes. But if this use of formative assessments becomes onerous, few teachers will continue using such assessments. State departments of education and school districts can lessen the burden on a state's teachers by creating assessment tools that can be used by those classroom teachers who wish to do so.

If the four recommendations I've trotted out in Section Two of this analysis were routinely implemented when educators assessed formatively, I believe this research-anointed assessment approach would be demonstrably more effective than if those recommendations were not followed. *Formative assessment* is not a magical talisman that, when used, will unerringly move students from performance that's wretched to performance that's rapturous. Formative assessments clearly can differ in their quality. But properly crafted and thoughtfully implemented formative assessments will, in fact, help students not only learn more, but learn it more deeply. We must allow our students to benefit from this powerful, instructionally focused assessment.

Postscript

As I noted at the outset, this paper was written primarily for the individuals who, in October, will be attending an Austin, Texas meeting to initiate the new CCSSO FAST SCASS. If others were to read the paper, I would not be lastingly offended. However, because the impetus for the paper was FAST SCASS, I'm snagging this opportunity to get my personal preferences on the table regarding what I hope this new SCASS will accomplish.

If the new FAST SCASS can substantially increase the number of American educators who (1) understand the differences between formative and interim assessments and (2) recognize how to make each of those assessments as effective as possible in improving students' learning, I will be elated. I realize there may be many more activities that might be undertaken by the FAST SCASS. But if we can only get our nation's educators to understand what a formative assessment actually is, and how it can be best used to benefit students, then I'll be more than happy to chalk this up as a major win for our FAST SCASS—and for the nation's children.

References

Black, Paul and Dylan Wiliam, (1998). "Inside the Black Box: Raising Standards Through Classroom Assessment," *Phi Delta Kappan*, 80, no. 2, 139-148.

Leahy, Siobhan, C. Lyon, M. Thompson, and D. Wiliam. (2005). "Classroom Assessment—Minute by Minute, Day by Day," *Educational Leadership*, 63, no. 3, 18-24.

Sadler, D. Royce (1989). "Formative Assessment and the Design of Instructional Systems," *Instructional Science*, 18, 119-144.

Scriven, Michael S. (1967), "The Methodology of Evaluation," in Ralph Tyler, Robert Gagne, and Michael Scriven, (Eds.) *Perspectives of Curriculum Evaluation*, (AERA Monograph Series on Curriculum Evaluation), Chicago: Rand McNally.