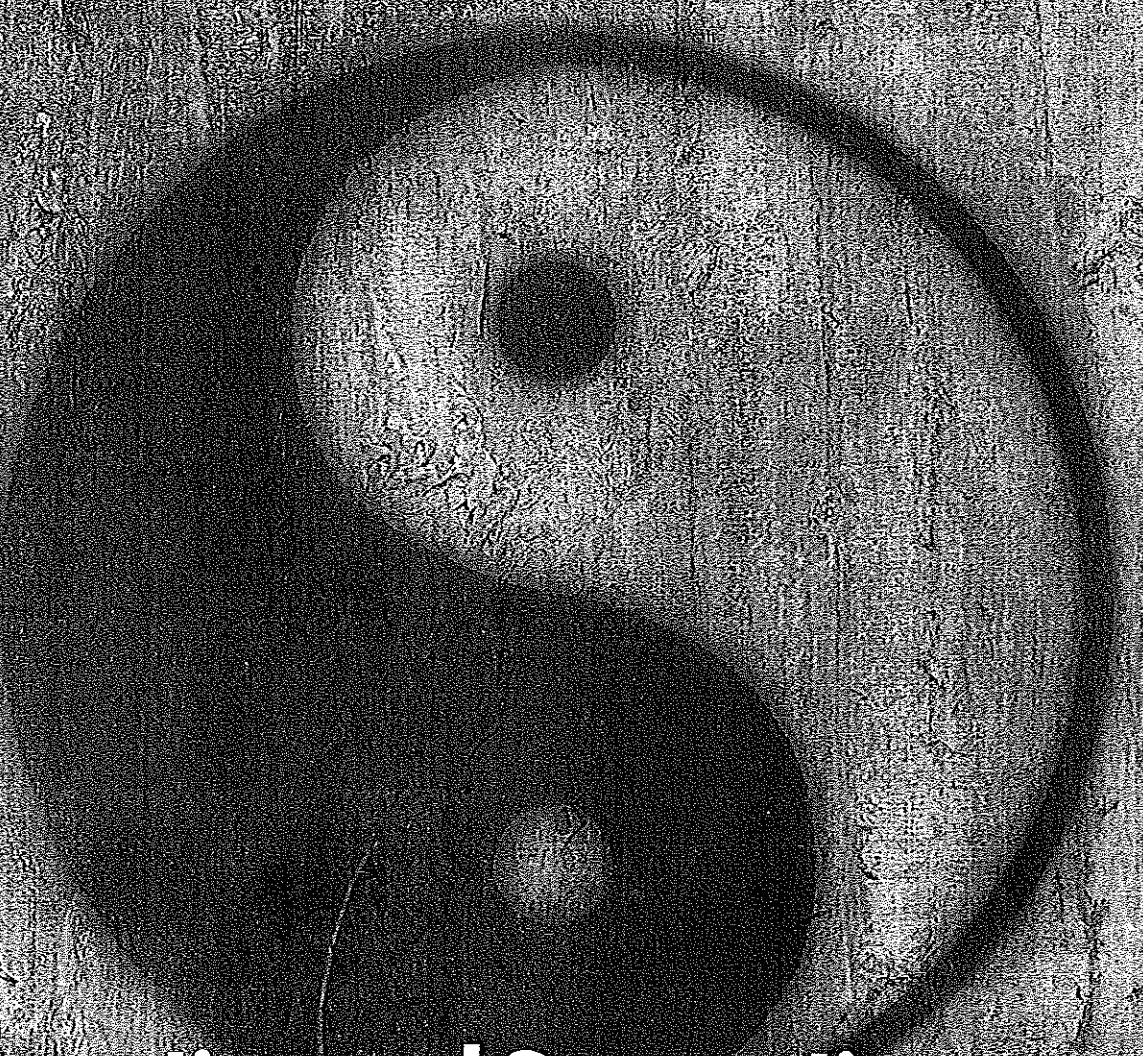




On Serving **TWO MASTERS**

**Formative and Summative
Teacher Evaluation**

W. James Popham

Although both formative and summative teacher evaluation are important, they must be completed separately to be useful and valid.

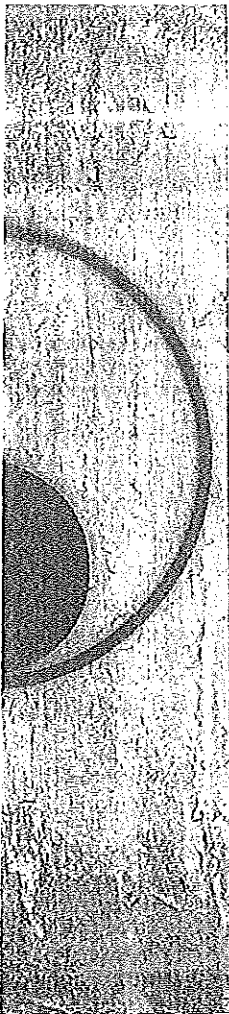
When I was growing up, my grandfather frequently tossed adages at me in an effort to make me a better person. Some were scriptural; some were secular. One of his favorites was “Waste not, want not.” My grandfather, you see, was thrifty. Another maxim that sometimes came my way was “No man can serve two masters.” Although this was well before the current era of gender equality, I am certain my grandfather was not implying that women were more adept than men at double-master serving. He just wanted me to recognize that simultaneously accomplishing two competing missions is sometimes impossible. This is precisely the dilemma currently facing the nation’s school administrators who are being asked to function—at the same time—as formative and summative teacher evaluators. It simply can’t be done.

A Pair of Federal Catalysts

Let’s consider what’s spurring so many of today’s school administrators to try serving, simultaneously, as formative and summative evaluators of their school’s teachers. As is far too fashionable these days, many blame this problem on the federal government. In this case, however, that accusation is altogether accurate. Two recent federal initiatives—the Race to the Top Program (RTT) and the ESEA Flexibility Program—have triggered a remarkable overhauling of the nation’s teacher-evaluation programs. All but a few states have substantially renovated their teacher-appraisal procedures in response to one or both of those federal initiatives.

In 2009, RTT provided substantial grants to cash-strapped states (and, later, to local school districts) if they established significant educational reform programs that incorporated rigorous teacher-evaluation procedures. Perhaps the most controversial feature of those recommended procedures was using student growth as a “significant factor” in determining a teacher’s quality. Not surprisingly, the allure of RTT dollars caused education policymakers in many states to meaningfully revamp their state’s teacher-evaluation procedures, sometimes statutorily, to increase the chance of receiving an RTT grant.

Two years later, in 2011, the ESEA Flexibility Program proved equally attractive to education officials in many states. That’s because states receiving a federal waiver under this program could successfully evade the increasingly harsh sanctions of NCLB—the most recent version of the Elementary and Secondary Education Act (ESEA). As was the case with RTT, for a state to successfully receive such a waiver, an alteration of the state’s teacher-evaluation system was required. The teacher-evaluation strategies recommended by both federal initiatives were essentially identical.



The distinction between formative and summative evaluation was first drawn by Michael Scriven (1967) in an influential essay regarding education evaluation.

Federal Guidance

To help state authorities as they attempted to revise their teacher-evaluation procedures, RTT and the ESEA Flexibility Program encouraged states to modify their teacher evaluations according to six specific guidelines (US Department of Education, 2012). To successfully compete for RTT cash or for ESEA waivers, a state's teacher-evaluation system needed to:

- Be used for continual improvement of instruction
- Differentiate performance on the basis of at least three levels
- Use multiple evidence sources to determine performance levels, including student growth as a significant factor
- Evaluate teachers and principals regularly
- Provide clear, timely, and useful feedback for professional development
- Be used to inform personnel decisions.

A perusal of those six guidelines reveals that some favor formative teacher evaluation and some favor summative teacher evaluation. Before considering those key distinctions, however, let's get more precise about what's meant by "formative evaluation" and "summative evaluation" when those terms are applied to teacher evaluation.

The Two Roles of Teacher Evaluation

The distinction between formative and summative evaluation was first drawn by Michael Scriven (1967) in an influential essay regarding education evaluation. Scriven supplied his analysis soon after ESEA was enacted in 1965—a time when almost no serious attention had been given to the nature of education evaluation. But because ESEA provided unprecedented federal dollars to support states' education programs and also required that the programs be evaluated, US educators soon clamored for guidance about how to do this thing called "education evaluation."

Scriven's analysis was both widely read and widely endorsed. He distinguished between the two roles of evaluation as they applied to the appraisal of education programs. Formative evaluation was intended to determine the worth—that is, the

merit—of not-yet-completed education programs.

The role of formative evaluation was to give educators information that could be used to improve still-malleable programs. Summative evaluation, in contrast, was intended to determine the worth of mature, already-completed education programs. The role of summative evaluation was to help decision-makers determine whether to continue or terminate a given program.

Although Scriven's distinction between the roles of formative and summative evaluation focused on education programs, the contrast between "improvement-focused" evaluation and "go-versus-no-go" evaluation was soon adopted by educators whose evaluative concerns dealt with teachers. Thus, on the basis of a distinction drawn almost a half century ago, what we refer to today as "formative teacher evaluation" is an endeavor intended to supply information that can improve a teacher's effectiveness. Summative teacher evaluation, however, provides information associated with negative decisions about a teacher, such as the denial of tenure and termination, or positive decisions, such as recognition and financial rewards. Formative teacher evaluation, then, is focused on improvement; summative teacher evaluation is focused on removal or rewards.

Looking back at the six previously listed teacher-evaluation guidelines supplied by federal officials, we see that two of them are clearly consonant with a formative conception of teacher evaluation. The first guideline indicates that teacher evaluation should be used for "continual improvement of instruction" and the fifth guideline calls for teacher evaluation to supply feedback for professional development. Both guidelines mesh well with a formative approach to teacher evaluation.

In contrast, however, the final guideline indicating that teacher evaluation should inform personnel decisions reflects a summative view of teacher evaluation. The other three guidelines, all important to be sure, are relevant to both formative and summative teacher evaluation. Whether one is a formative or summative teacher evaluator, it makes good sense to evaluate regularly, employ multiple sources of evidence, and differentiate teachers' performance using different levels.

As state-level architects of recently fashioned teacher-evaluation systems set about to devise evaluative procedures that would pass federal muster, it is not surprising that many of those systems attempted to blend the formative and summative roles of teacher evaluation. After all, federal authorities seemed to suggest that they wanted an approach capable of simultaneously supplying information for both formative and summative purposes. Regrettably, dual-mission teacher evaluations are destined to fail.

An Evaluative Dilemma

The reason that dual-mission teacher evaluation won't work resides in human nature. Teachers want to improve their skills. I've never met one who didn't. But teachers also want to keep their jobs. I've never met one who didn't. Realistically, with few exceptions, job-keeping trumps skill-improving.

Most teachers are willing to reveal their instructional shortcomings to an evaluator whose exclusive function is to help the teacher get better. Teachers realize that their instruction deficits can be addressed only when they have been identified. But those very same teachers will be reluctant to identify their instructional weaknesses when those shortcomings might result in their dismissal.

An evaluator whose responsibilities include even a dollop of summative information gathering might as well abandon any hope of getting a teacher to forthrightly reveal his or her instructional flaws. The potentially adverse consequences of summative teacher evaluation invariably contaminate the efforts of formative teacher evaluators.

But the teacher evaluation coin has two sides, and although summative teacher evaluation can mess up formative teacher evaluation, it is equally true that formative teacher evaluation can diminish summative teacher evaluation. Let's illustrate that point using an example in which a summative teacher evaluator employs classroom observations to help determine a teacher's skill. For classroom observations to yield the most valid estimate regarding how well a teacher instructs a class, those observations must be unannounced; prescheduled classroom observations tend to result in staged, atypical pedagogical performances. Yet we find many school

administrators who resist the use of unannounced classroom visitations because such observations might reduce the collegial rapport those administrators believe necessary to support a collaborative, more formative approach to teacher evaluation.

"Hybrid teacher evaluators" attempt to carry out both formative and summative functions. But mixing those often-opposing evaluative missions simply doesn't work. The frailties of human nature, not the least of which we can see in a teacher's self-preservation behaviors, make hybrid teacher evaluation more of a wish than a sensible evaluative strategy.

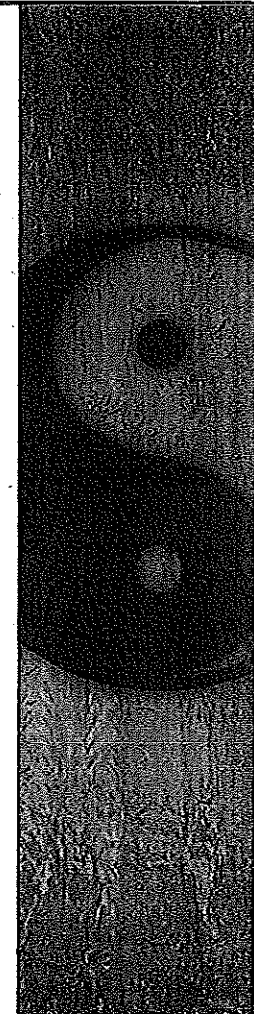
A Solution

Clearly, both the formative and summative missions of teacher evaluation are tremendously important. But if they contaminate each other, then the obvious solution is to engage in both of them, but to do so separately. That is, evaluators must distinguish between formative and summative teacher evaluation and keep those two roles apart—totally. For example, suppose a high school principal serves as a school's formative teacher evaluator, and the assistant principal tackles the needed summative evaluation chores. Reassurances must be given to teachers, of course, that no information from the formative evaluations will ever be relayed (formally or informally) to the summative evaluator. This is another instance where "divide and conquer" wins the day. Separating formative and summative teacher evaluation helps school administrators attain the considerable benefits of both of those evaluative endeavors.

Some school administrators will protest that they can establish a suitable level of trust with their school's teachers so that both formative and summative teacher evaluation can be carried out by the same individual. Over the years, I've spoken with more than a few administrators who harbor such a hope. But although I understand why it is that they want to believe they can pull off this sort of schizoid undertaking with success, they are almost always fooling themselves. Human nature, with all of its fundamental foibles, always triumphs.

A Backward Look

Twenty-five years ago I wrote an article entitled



The reason that dual-mission teacher evaluation won't work resides in human nature. Teachers want to improve their skills. I've never met one who didn't. But teachers also want to keep their jobs.

"The Dysfunctional Marriage of Formative and Summative Teacher Evaluation." The article's opening sentence revealed my disdain for the teacher-evaluation practices of that era: "The evaluation of America's schoolteachers is, with few exceptions, an anemic and impotent enterprise—promising much but producing little" (Popham, 1988, p. 269).

In that essay, I railed against the misguided attempt of US teacher evaluators to navigate a middle course between formative and summative evaluation. I've not altered my position on this issue—not even a little. I suppose my grandfather could have invoked some sort of adage dealing with the virtues of "staying the course," but he never tossed that particular maxim at me.

What we must realize today is that 25 years ago, teacher evaluation was an ineffectual undertaking, an undertaking doing little good—but doing little harm as well. Now we see new, high-stakes teacher evaluation systems emerging all over the United States. They can do plenty of good, but they can also do plenty of harm. Inappropriate evaluation will have a disastrous impact on schools. The unthink-

ing mixture of formative and summative teacher evaluation will almost certainly foster inappropriate teacher evaluations. But because formative and summative teacher evaluation can each make significant contributions to instruction, both should be implemented widely—but separately—as components of emerging teacher-evaluation systems. **PL**

REFERENCES

- Popham, W. J. (1988). The dysfunctional marriage of formative and summative teacher evaluation. *Journal of Personnel Evaluation in Education*, 1, 269–273.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. W. Gagne, & M. Scriven (Eds.), *American Educational Research Association Monograph Series on Curriculum Evaluation: Vol. 1. Perspectives of curriculum evaluation*. Chicago, IL: Rand McNally.
- US Department of Education. (2012). *ESEA flexibility*. Retrieved from www.ed.gov/esea/flexibility

W. James Popham is an emeritus professor in the Graduate School of Education and Information Studies at the University of California–Los Angeles. His book *Evaluating America's Schools: Mission Possible?* is forthcoming from Corwin Press.



"The knowledge and skills acquired at Seton Hall and the network of cohort colleagues were essential elements to my career advancement."

Jason E. Glass, Ed.D. '11
Iowa's State Director (Commissioner) of Education

A Fast Track Traditional Ed.D.

WHY MORE ADMINISTRATORS HAVE CHOSEN Seton Hall University's Executive Ed.D. Program

- National/International Reputation
- Intensive 2-year program: 10 Weekends, Two 4 week Summer sessions
- Cohort Model of 30 Students
- Dissertation Begins Day One
- Personal Care and Support

Now Accepting Applications for Cohort XVII

For more information, call 1-800-313-9833, email execedd@shu.edu, or go to www.shu.edu/go/execedd.



COLLEGE OF EDUCATION
AND HUMAN SERVICES

SETON HALL UNIVERSITY

400 South Orange Ave. • South Orange, NJ 07079