

The Role of Student Achievement in the Evaluation of Teachers

Rick Stiggins Assessment Consultant Portland, Oregon

October 2012

“In exchange for absolution from meeting No Child Left Behind achievement expectations, educational leaders across the nation have agreed, among other things, to implement a teacher evaluation system that factors evidence of student achievement into the evaluation process.”

In exchange for absolution from meeting No Child Left Behind achievement expectations, educational leaders across the nation have agreed, among other things, to implement a teacher evaluation system that factors evidence of student achievement into the evaluation process. I believe student achievement data can play an important role in the evaluation of teachers, but only if certain essential conditions are satisfied. Given the history of policy maker thinking about this matter, I am certain educators across the state are thinking that this merely codifies the expectation that teachers will be required to demonstrate gains in student performance on annual accountability tests. However, at this time and for several very specific reasons, such tests cannot serve as the evidentiary source for the evaluation of the instructional impact of individual teachers. But the good news is, there is another approach to the consideration of student achievement data that can work, again, under certain conditions.

In order for a once-a-year test to serve as the basis for judging the performance of an individual classroom teacher, it would need to provide evidence of student mastery of all of the learning targets that teacher is expected to impart during the period over which the evaluation is conducted. In other words, consideration of only math and reading scores would provide an inadequate sample of teacher impact when a teacher is responsible for four or five additional preparations such as science, social studies, or writing. In effect, such a thin sample would ignore key learning targets on which the teacher may have, in fact, performed well or poorly. That would be unfair. Total coverage is essential. What annual scores would be used to evaluate the performance of those middle school and high school teachers who don't teach math or reading? What scores would be used across grade levels to reflect changes in the achievement of special needs students. What cores would be used to reflect those 21st Century proficiencies needed to success in the digital age that cannot be reflected on the multiple choice tests currently in use?

In addition, annual accountability tests typically report only undifferentiated total scores, thus once again masking the strength or weakness of a teacher's performance. Evidence used to judge the performance of individual teachers would need to inform evaluators about how each student did in mastering each relevant achievement standard in order to judge its success. Precise evidence is essential. There are no tests currently under development in state or national pipelines that report evidence in this form for each teacher at each grade level in each relevant subject.

To serve in the capacity of detecting the impact of an individual teacher, an assessment would need to be shown sensitive enough to detect the

effects of different levels of teacher's instructional skills on student achievement. Only then could an assessment reveal a teacher's true impact on students' learning. Until instructional sensitivity is demonstrably verified for an assessment, it cannot serve in this capacity.

Finally, if teachers are to be judged and compensated based on achievement evidence, the assessment and the evaluation process would need to be conducted in such a manner as to minimize rival explanations for the achievement levels revealed. In other words, school leaders need to come as close as possible to establishing a causal link to learning effects. Rival explanations for those effects, such as factors that are beyond the control of the teacher, would need to be systematically eliminated. Analyses that merely demonstrate the co-occurrence of some level of student achievement and the presence of a teacher in that classroom, such as currently popular value-added analyses, fall short of what is required for a dependable judgment if a teacher's impact on learning.

Local districts do not have now nor has it ever had at their disposal assessments that provide sufficient coverage and precision of evidence or demonstrated instructional sensitivity. Further, there is no pipeline for such assessments in the future.

But this does not mean schools can't consider student achievement as a significant factor in the evaluation of teacher performance. There is an alternative source of evidence could be brought to bear. We could rely on evidence generated with classroom assessments. This evaluation process is simple to describe, but would be demanding to implement. If successfully implemented, the benefits will be legion.

Within this system, each teacher would begin the instructional year with a specific set of achievement standards her or his students are to meet. Each teacher can be charged with developing and conducting an assessment program that would document the extent of student attainment of each of those standards. At the end of the evaluation period, teachers would be required to prepare and present an annual report to their employer detailing that evidence, defending in the strongest possible terms (a) the quality of the evidence they offer, and (b) their case for having aided students learning. Their reports could include, but would not be limited to, evidence from annual tests or other externally produced test if that evidence is reflective of student mastery of an assigned achievement standard.

Evidence also could be mustered from the classroom assessments conducted during the school year. Commentary on teacher effectiveness as reflected in that evidence would be supplied by teachers themselves. Instructional supervisors for the school and district would then review and evaluate each teacher's presentation, judging performance based at least in part on the evidence of student learning presented.

This process addresses each of the criteria identified above. Comprehensive coverage of the teacher's content area responsibilities could be address with proper sampling of available assessment evidence. That evidence could be gathered and reported with appropriate precision (i.e., by standard), by using proper sampling procedures. The instructional sensitivity of assessments used could be enhanced by relying on assessment tasks linked directly to the teacher's own specific instruction.

"..if teachers are to be judged and compensated based on achievement evidence, the assessment and the evaluation process would need to be conducted in such a manner as to minimize rival explanations for the achievement levels revealed."

"Finally, the impact on student learning of any factors that are beyond the control of the teacher could be addressed by the teacher in the final report."

And rival explanations for effects could be minimized, again, because of that tight link between assessments conducted and reported to a single teacher's instruction. Finally, the impact on student learning of any factors that are beyond the control of the teacher could be addressed by the teacher in the final report.

But, as mentioned above, certain essential conditions that would need to be satisfied for this to work effectively. It is worthy of note that these four conditions have rarely been satisfied in schools and districts across the county:

- Achievement standards would need to be articulated in all relevant subjects or courses for each teacher, and they would need to be agreed to by all involved at the beginning of the year
- Teachers would have to be provided with the professional development needed to attain sufficient levels of assessment literacy needed to gather, summarize, report, and defend their evidence of student learning in their classrooms
- The evidence presented of mastery of the standards would, in fact, need to be of verifiably high quality; teachers would need to build that into their final report and supervisors would need to be qualified to evaluate that part of the presentation
- Therefore, those in supervisory roles who evaluate each teacher's case also would to be given the opportunity to develop sufficient levels of assessment literacy needed to be able to judge the quality of the evidence presented by the teacher and the strength of the case advanced by each teachers of their instructional impact

It is worthy of repeating for emphasis that this system would require attainment of levels of assessment literacy throughout the system that have typically not been made available to the staff and faculty of our schools.

But, if these conditions are satisfied, students would benefit from profoundly important side benefits. For example, the resulting foundation of assessment literacy would permit far stronger links among quality classroom assessment, improved day-to-day instructional decision making, and student learning success than are in place today. Used well, those links have been shown in research conducted around the world to promote impressive gains in student learning, especially for struggling learners. The vast majority of practice teachers are not trained to use these new classroom assessment interventions. Another benefit would be the positive statement such a system would make about the professionalism of practicing teachers. While objective third party evidence in the form of annual test evidence could be brought to bear if relevant, so too would evidence of achievement gathered by our professionals themselves in their classrooms.