

Don't go there [yet]!



Instructional Support blog post.

Now that most of the SLO-setting work has been completed, Lead Evaluators in schools across the state are stepping up their classroom evidence gathering. In some cases this is still done through the traditional preconference/observation/postconference process. Increasingly, however, more frequent mini-observations are taking the place of the "dog and pony shows" of the past. No matter the format, however, evidence is being collected and shared with teachers according to the spirit of the new Annual Professional Performance Review (APPR) process in New York.

The basic idea is that evidence is collected by the Lead Evaluator. That evidence is then sorted according to the New York State Teaching Standards using a framework or rubric. Once the evidence has been collected and sorted, it is shared with the teacher and a conversation about that evidence occurs. The best conversations will be growth-producing conversations in which both evaluator and practitioner reflect upon the collected evidence and consider its implications on future practice. The entire process is repeated some number of times across the span of the school year. In the process legislated for New York State, a summative evaluation is also conducted at which time the teacher is given an overall score for the year that is based on some measures of student achievement and a number of points derived from the rubric. That summative score satisfies the requirements of the legislation and accompanying regulations. It's up to local school districts to implement the process in a way that is about growth-producing feedback and continuous improvement rather than the inauthentic system of fear and inspection that Deming warned us about.

Some districts, however, are making a key mistake, early in the process, that will hamper the process and negatively impact the possibility of the system resulting in continuous improvement, better teaching, and increased student achievement. The mistake that some districts are making might seem like a subtle one but the consequences can be significant. Instead of waiting until the end of the year and using an accumulated collection of evidence from throughout the year that is based on many classroom visits and evidence submissions, they are rating the teacher on the rubric *after each episode of evidence collection*. This is a mistake and contorts the whole system to be about judgement and labelling rather than growth-producing feedback and continuous improvement.

We are all well aware of the shortcomings of the legislation and the accompanying rules for implementation -- rules that were crafted without the context of the field being taken into consideration (20+20+60, for example). Despite that, we *can* make the best of the situation and make it about improving the teaching and learning process. When we rate teachers every time we see them, however, we are just continuing the old paradigm of judgement rather than shifting to the new, better paradigm of continuous improvement. The experience of districts who have rated teachers on the rubric after each observation unfortunately confirms this. If you rate teachers after every evidence collection the conversations that follow inevitably become conversations about the score rather than conversations about the teaching and learning.

There's another reason for not scoring the rubric after every episode of evidence collection: *the rubrics just don't work that way*. A close read of the NYSUT rubric identifies stems such as: "Teacher provides regular opportunities..." and "Teacher frequently uses..." A close read of the Framework for Teaching (2011) identifies stems such as: "Assessment is regularly..." and "All outcomes represent..." Based on the way these, and other, indicators are phrased it is simply impossible to accurately rate them based on a single episode of evidence collection. To try and do so will result in a lower rating than might otherwise be the case with evidence collected over multiple opportunities. This cannot be avoided because of the way the rubrics are written. It is unfair to rate after every single episode of evidence collection – and fairness is one of the three gates (in addition to reliability and validity) that must remain open if a system of teacher evaluation is going to have a shot at making a difference.

The remedy is simple: don't go there! Don't rate teachers until the end of the year when the preponderance of evidence can be compared to the levels on the rubric. It's still early in our first year of implementation of the system and unless your agreed-upon system clearly requires scoring after every evidence collection episode you can stop the practice immediately and chalk it up to a learning process in which we are all engaged.



**DON'T
EVEN
GO THERE**

Stop judging after every visit. Instead, make it about a process of growth-producing feedback that is designed to continuously improve teaching and increase student learning. That is, after all, what it's supposed to be about.