

Questions and Answers

Implementing a Comprehensive Instructional Effectiveness Assessment System

K-12 Webinar March 26, 2013, Hosted by Assessment Technology Incorporated

Presented by Jason K. Feld, Ph.D., Vice President Corporate Projects

Q: Can a district set different weightings for staff? For example, can weightings for teachers of state-tested content be different than for teachers of non-state-tested content.

A: Yes. A district can in fact construct a different template including a separate set of categories for each group of teachers being evaluated (e.g., teachers of state-tested content versus non-state-tested content). Each category within each template can be weighted separately with the constraint that the sum of the weights cannot exceed 100 percent. The goal here is to make the system as flexible as possible so that districts have a large measure of control over it. Once that's done, then you're ready to go.

Q: Can we incorporate other scores into the system? Can we put school performance scores in or surveys into the weighting?

A: Yes. External scores can be entered into the system and assigned to a category in a template. At that point, a weight can be assigned to the category in the same way that weights can be assigned to any other category. The scores and weights for all of the categories in the selected template will appear on the *Evaluation Score Compiler*, which provides an overall evaluation score for each educator assigned to the selected template. Additionally, the *Evaluation Score Compiler* allows you to perform test runs. That is, you can put weights and categories into the system and then actually do some research locally to see what types of outcomes you get in terms of overall effectiveness when you look at different weights and different categories.

Q: Our new teachers will require more observations throughout the year than an experienced teacher. Can we monitor the groups separately?

A: Yes. Within the *Dashboard* you can actually set up different templates with different requirements, for different groups of teachers. The number of observations that you care to do during a school year for any set of teachers is really totally up to you. What we're doing is providing you with an easy and convenient interface so that you can see a calendar for those observations, you can have a history of those observations, and then throughout the year, you can actually generate an ongoing effectiveness score so that, as we often say, you don't have to wait until the horse is out of the barn at the end of the year to actually do good things to make things work better within your school district.

Q: Can we use existing benchmark assessments or teacher-constructed assessments for measuring student progress?

A: Yes. Typically, in a systematic approach to educator effectiveness initiatives, a district is recognizing that they're simply not collecting data to evaluate a teacher or a school leader. The real purpose in gathering data is to do something - to take an action. The action can be to identify the kind of professional development that's needed, it could be action related to what needs to be

changed in the classroom to move students from higher to lower levels of risk in not meeting Common Core State Standards or to motivate higher achiever, lower growth students. And if we conceptualize an educator effectiveness initiative as one that is designed to assist teachers and leaders in enhancing their impact on student learning, then reliable, valid, and timely data for use in decision-making becomes really important.

To support assessment in areas not covered on statewide assessments, ATI has launched the *Community Assessment and Item Banking Initiative*. We're hearing that districts would like to get more engaged in building assessment items and tests in non-state-tested areas. The *Community Assessment and Item Banking Initiative* makes it possible for districts to not only build, review, and certify items in these non-state-tested areas, but also to share and receive item content as a part of a state-wide and multi-state learning community.

Q: Which measures of student progress are recommended for non-state-tested subjects?

A: At the end of the day, the local school district will want to come together as a community to make decisions regarding what measures of student progress will be used to measure growth. In some states that we're working with, there's a strong push to have teachers that are in non-state-tested areas develop and use reliable and valid tools to measure student growth in those non-state-tested areas. That's one of the reasons we began the community item banking initiative awhile back. The capacity of the *Evaluation Score Compiler* within Galileo® to accommodate any category and set any weight to it, is functional regardless of what a district decides to do with regard to measurement of student growth.

Q: Are the pretest and the posttest identical, or do they follow the same blueprint?

A: The pretest and the posttest are not identical, although they are both designed to be comprehensive in nature. The reason they are comprehensive is because we want to get as broad and deep a view of student learning over the course of the school year as we possibly can. The pretest and posttest are designed to do that. The pretest does in fact include standards and items from the prior school year – so if you are a fourth-grade student and you are in a classroom where the teacher is administering the pretest, there will be some items from standards that would appear in the third grade. We do that for a number of reasons. For example, it is obvious that if you assess students on everything they may not have learned, you are not going to get much variation in student performance. As a result, the assessment is not going to be as valuable as you would like in getting a good baseline of what the students know so that you can use the information to help guide instructional decision-making. The posttest on the other hand is designed to cover in a comprehensive way everything that was identified as standards required for students to learn within a school year. We do ensure that there is a sufficient overlap of standards between the pretest and posttest so that we can place the scores on a common scale using Item Response Theory techniques, and generate meaningful information about growth. Use of raw score data will not accomplish this goal.

Q: I used the Categorical Growth Summary/Student Growth and Achievement Report to evaluate class growth. Do I compare the pretest with the latest benchmark, or Benchmark 1 with the latest benchmark?

A: With a Categorical Growth Analysis you can compare pretest data with any Galileo benchmark assessment administered throughout the year in order to obtain information about student growth and achievement throughout the year and use that information to guide instructional practices.

Q: How did you decide which graphs to show in dashboards? Which parts of the dashboards have districts found most helpful?

A: The graphs that appear in the dashboards are intended to provide teachers, school and district administrators with easy access to information that is actionable and can be used to inform decision-making at all levels of the district and through a common “data language” so that communication is meaningful and purposeful. Most districts find all *Dashboard* components useful.

Q: Can teacher standards be customized if our district is not using the standards in the demo?

A: Yes, the item banking tools within the Galileo *Dashboard* technology make it possible for districts to build and customize educator proficiency scales that meet their local needs.

Q: In the non-state-tested content areas, would we create a pre/posttest and would this then be used in the Categorical Growth Analysis?

A: Yes. We can work with districts to generate Categorical Growth Analyses for non-state-tested areas.

Q: Is this an evaluation tool or observation tool? My biggest issue is that there are so many things to monitor, I usually have a list for some things, but write an ongoing narrative to refer back to. How does this tool help me with that?

A: The system is designed to make it possible for both to occur. As discussed during the webinar, Galileo® item banking technology makes it possible to: 1) construct locally customized rating scales aligned with INTASC and ISLLC standards; 2) place local, district-developed scales into Galileo and modify them as needed; 3) place third-party scales into Galileo so long as the district has copyright permission to do so; 4) import data from scales administered outside of Galileo into the system; 5) link educator performance criteria to Common Core State Standards requirements; and 6) accommodate changes in standards and assessment requirements over time. As a result, districts can then use Galileo to accommodate scheduling, security, data entry and analysis, and reporting capabilities. Moreover, scale administration process provides easy access to: 1) standards and performance competencies being observed; 2) scoring rubrics to guide performance ratings; 3) data sources documenting performance ratings; 4) anecdotal notes and calendar information; and 5) integration with educator portfolio and other supportive materials.

Q: Is it possible to create an option in the Categorical Growth Report to allow the report to disaggregate data looking at special education, gifted, or English language learners?

A: ATI is looking into the best ways to provide this type of information to educators.

Q: Will the Categorical Summary Report be available for download in Microsoft Excel?

A: ATI is planning to include this functionality in the near future.