

Improving and Validating Teacher Evaluation Systems

Empirical is an eight-year old research company located in Palo Alto, CA. The company has a strong team of 25 researchers, statisticians, engineers, and managers. Our basic mission is to help school districts obtain the evidence they need to make productive and cost effective decisions about adopting, continuing, and improving products, programs, and policies. Our goal is to support districts in answering their questions through careful examination of the results of their local program implementations and appropriate use of advanced analytic methods. Rather than serve only as an external evaluator, we see our role as supplementing and amplifying the capabilities for research, including collection of all types of data, compilation and analysis of data, and reporting in a format that is accessible to practitioners, administrators, and board members.

Validation Engine

Empirical Education is currently developing a web-delivered software tool called the Validation Engine that allows educational agencies to analyze and validate the observation protocols they use for teacher evaluation observations. The tool is being developed with funding from the Bill & Melinda Gates Foundation and builds on the large corpus of classroom video and student data collected as part of the Measures of Effective Teaching project. This product will be useful to any school system that is developing an observational rating system. We are currently engaged in beta tests and invite participation in this activity.

Whether an observational protocol is homegrown or adapted from popular protocols, it is important to test whether at least some elements of the protocol award the highest ratings to teachers whose students' academic performance grew most and their lowest ratings to teachers whose students' academic performance grew least. The Validation Engine validates these protocols against one or more standardized observational or analytical benchmarks (calculated with value-added modeling), and gives users information about the reliability and consistency of their "raters" (i.e. staff who would be utilizing these protocols in actual classroom observations).

School systems use the Validation Engine to view a collection of classroom observation videos, rate those videos with their own observation protocol, and then receive a report that evaluates the predictive validity and rater consistency for the protocol.

The Validation Engine has four components:

1. A **secure video viewer** that provides web-based access to a video library of teacher lessons for raters to watch.
2. A **scoring template** that allows raters to enter scores for each observed video into a web-based system on each domain of the observation protocol.
3. An **analytic engine** that includes value-added benchmarks on state and supplemental assessments.
4. An **automated report generator** that provides feedback on the predictive validity of the observation protocol as scored how it performs against a set of expert-scored benchmark instruments.

Process

Teacher evaluation systems should be evaluated and refined on a continual basis. The Validation Engine (VE) can assist school systems in a piece of their comprehensive teacher evaluation system, specifically, their classroom observation protocols. Objective measures of student achievement gains, student and teacher perceptions, teacher reports/check-ins, and other such methods should also be factored in when compiling a comprehensive teacher evaluation system.

The Validation Engine provides information to assist districts and school systems to improve and update their observation protocols by identifying certain areas that don't correlate with established benchmarks, and work on obtaining high inter-rater reliability among their classroom observers by identifying raters who may need more training than others.

The VE assists in measuring the reliability of the observational protocol through its training and inter-rater reliability functions. The validity of the VAM scores themselves are based on extensive statistical modeling conducted by third parties who are leaders in VAM theory. The VE considers the potential measurement error in constructing the VAM from the student data and considers that in its evaluation of the observational protocol.

Technology

Since the Validation Engine is a web-based service, no installation of product software is required to use the system. School systems will need access to the observation protocol they wish to evaluate and users will fill out the protocols while viewing the VE-supplied videos the same way they fill out the protocol when conducting a live observation.

The VE allows users to conduct an unlimited number of studies, and the automatic report generator is able to produce multiple interim and final reports about the validity of the protocol and the consistency of their protocol testers/users. The reports are downloadable and always accessible on the Validation Engine website.

Timeline

Conducting an evaluation study of an observation protocol through the Validation Engine does require the availability of multiple school system staff. Depending on the number of staff available and the time they have to devote to the project, school systems could complete an evaluation of their observation protocol in a matter of weeks. The automated report generator provides instantaneous results that can inform further development and revision of the protocol, as well as identify raters (protocol users) who may be in need of further training.

The identification of these areas of need and improvement may help school systems better focus their professional development efforts, saving time in the long run.

The Validation Engine is still under development and slated for delivery and launch later this year.

For more information, please contact:

Kylene Chinsio Shen
Product Manager
T: (650) 328-1734 x137
kshen@empiricaeducation.com
www.empiricaeducation.com