**Teacher Effectiveness: Can We Measure this in an Informative Way?**

Franklin W. Bender
University of Oregon

---

### Background

- The Obama administration has suggested that an updated and educational policy and accountability system could be our “moonshot” of the 21st century to ensure that the United States is economically competitive within the international community.1
- Fowler (2009) suggests that the nature of reform will be imposed by the presence of a problem where constituencies have divergent values. With the reauthorization of No Child Left Behind (NCLB) currently being discussed at the federal level, the educational accountability system, vertically spanning the federal, state, and local levels, is at a precipice where operational changes need to be considered.2
- One of the educational discussions points being seriously investigated in the educational policy arena is how educators should be objectively evaluated and this data can be used to potentially predict student educational performance.

### Defining Teaching Effectiveness & Noting the Problem

- The general consensus within the literature is that teachers play a critical role in the performance of students. However, there are differences in defining and measuring the the construct associated with what is referred to as “teacher effectiveness”. Teacher effectiveness will be defined here as a teacher’s ability to improve student learning as measured by summative assessments.3
- There is a paucity of empirical research reflecting predictive and causal relations between teacher effectiveness and student academic outcomes due to the inherent validity and reliability constraints associated current models associated with measuring teacher effectiveness.4
- The paradox of measuring teaching effectiveness is that student assessment results appear to be the best proxies for measuring teacher effectiveness. Teachers are the most proximal and malleable variable that can be controlled to impact student outcomes. However, covariates that are more distal and rigid (e.g., socioeconomic status and family involvement) impact the discussion of this topic between policy actors.5

### Recommendation: Comprehensive Measurement Tool

- The inherent challenges associated with measuring teacher effectiveness is that it encompasses multiple constructs nested within latent constructs (e.g., a variable that can’t be measured directly).6
- Because of this, a triangulated approach to collecting and interpreting multiple teaching constructs should be used to measure teaching effectiveness. This type of comprehensive evaluation would embody a weighted matrix that would yield a relevant composite score that captures the attributes of effective teaching.7

### The MET Project

In the fall of 2009, the Bill and Melinda Gates Foundation launched the Measures of Effective Teaching (MET) project in order to develop and test multiple measures of teacher effectiveness. The goal of the project is to build fair and reliable systems for teacher observation that can be used for a variety of purposes, including instructor feedback, professional development, and continuous improvement that will address educational accountability requirements. Even through the MET Project recognizes that student achievement gains need to be valid predictors of teacher effectiveness, gain scores alone cannot be the sole determinant of capturing and evaluating teacher effectiveness. The five areas studied were: (a) student achievement, (b) classroom observations and teacher reflections, (c) teachers’ pedagogical content knowledge, (d) student perceptions of the classroom instructional environment, and (e) teachers’ perceptions of working conditions and instructional support at their schools.8

### Methods & Results

- Large scale study randomly assigning 1,300 teachers to classroom from school districts from 6 different states.1
- 5 Observation Instruments were used (e.g., Framework for Teaching – FTF; Classroom Assessment Scoring System – CLASS; Protocol for Language Arts Teaching Observations – PLATO; Mathematical quality of Instruction – MQI; Utah Teacher Observation Protocol – UTOP)9
- 7,491 videos of lessons were scored at least 3 times by raters.10
- Test scores and student surveys were incorporated from 44,300 students.10

### Multiple Observations Lead to Higher Reliability

#### Testing Validity

To What Extent Do Indicators Predict Outcomes?

**Teacher Indicators**

<table>
<thead>
<tr>
<th>Teaching Quality</th>
<th>To What Extent Do Indicators Predict Outcomes?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
<tr>
<td><strong>Teaching Quality</strong></td>
<td><strong>To What Extent Do Indicators Predict Outcomes?</strong></td>
</tr>
</tbody>
</table>

---

### References


### Methods & Results

- Large scale study randomly assigning 1,300 teachers to classroom from school districts from 6 different states.1
- 5 Observation Instruments were used (e.g., Framework for Teaching – FTF; Classroom Assessment Scoring System – CLASS; Protocol for Language Arts Teaching Observations – PLATO; Mathematical quality of Instruction – MQI; Utah Teacher Observation Protocol – UTOP)9
- 7,491 videos of lessons were scored at least 3 times by raters.10
- Test scores and student surveys were incorporated from 44,300 students.10

### 3 Key Take Ways from MET Project

- High-quality classroom observations will require clear standards, certified raters, and multiple observations per teacher. Resources allocated for observation training, purchase of evaluation tools, and continued professional development will be essential.1
- Combining the 3 approaches (classroom observations, student feedback, and value-added student achievement gains) capitalizes on identifying teacher strengths and offsetting weaknesses.9
- Combining new approaches to measuring effective teaching (while not yet perfect) significantly outperforms traditional unidimensional measures. By providing more valid and reliable evidence, better educational decisions can be made.10

### Contact Information and Acknowledgements

For further information, please contact Franklin W. Bender, University of Oregon: fbender@uoregon.edu.