Using Classroom Observation to Gauge Teacher Effectiveness:
Classroom Assessment Scoring System (CLASS)

Bridget K. Hamre, Ph.D.

Center for Advanced Study of Teaching and Learning
University of Virginia
Overview of Presentation

- Classroom observation as one tool for measuring and improving teacher effectiveness
- Factors to consider when choosing an observational tool
- CLASS Overview and Research
- Using observation to improve teaching at scale - examples from the field
Goal: An aligned system of teacher evaluations and supports

Teacher Evaluation and Support System
- Evaluation
  - Principal Evaluation
  - Observation
  - Tenure and Incentives
  - Hiring
- Support
  - Professional Development
  - Teacher Peer Networks

Teacher Effectiveness
- Student Achievement Data
- Effective Teacher-Student Interactions Data (CLASS)

Administrative Support

Cross Grade/Content Alignment

School and District Alignment

Teacher Buy-In
CLASS Development Process – 15 Years

- **Stage 1 – What matters?** Conceptual work and basic research on elements of teaching practice that directly contribute to student learning and development
- **Stage 2 – How do we measure it?** Development and field testing of observational tool – focus on balancing reliability and validity
- **Stage 3 – Does it work?** Rigorous evaluations of validity
- **Stage 4 - Can we change it (and if we do does it improve student outcomes)?** Development and testing of professional development to support improvements in these practices
- **Stage 6: Can we implement at scale?** Development of structures to support use of tool and aligned professional development at scale
Choosing an Observational Tool
Choosing an Observation Tool

- **Scope and Alignment**: Is the scope of the tool aligned with the types of outcomes of interest?

- **Standardization**: Are standardized observation and scoring protocols available (manuals, scoring sheets, etc)?

- **Training**: Is there adequate training available for use at scale (e.g. Train the Trainer)?

- **Reliability**: Is there evidence that observers can use the tool reliably at scale?

- **Validity**: Is there a credible research base linking scores from the tool to outcomes of interest in our population?
Our Goal – high-inference but well tested measures

Current system of teacher evaluation in most districts

Cannot have high validity with low reliability

Low inference, standardized behavioral measure
Other considerations

- **Feasibility**: Is the time required for training and observation feasible for your organization? Are the requirements for who can observe reasonable?

- **Sensitivity to Change**: Does the protocol have evidence that it is sensitive to assessing change (e.g., based on intervention such as participation in professional development)?

- **Links to Improvement**: Does the observation include guidelines and support for using findings for professional development purposes?
Overview of CLASS
Scope and Alignment
Student-teacher interactions and schools

- Interactions with adults form “infrastructure” for school success:
  - Cognitive processes, language
  - Self-regulation, emotional self-control
  - Task orientation, persistence, motivation, engagement

- Instruction is, in part, a social process:
  - Interactions with teachers are a “medium”
  - Good “instruction” is embedded in relationships and interactions

- Interactions operate across content
What is the CLASS?

• CLASS is a tool for observing and assessing the effectiveness of interactions among teachers and students in classrooms.

• It measures the *emotional, organizational, and instructional supports* provided by teachers that have contribute to children’s *social, developmental, and academic achievement*.

• CLASS is used to assess interactions among teachers and students for a variety of purposes:
  – Teacher Professional Development
  – Monitoring and Evaluation of Teacher Performance/Effectiveness
  – Research
CLASS versions – Coherence, systemic

- Infant (CLASS-I) – in validation
- Toddler (CLASS-T) – fully supported
- Pre-Kindergarten (CLASS-Pre-K) – deployed at scale
- Elementary (CLASS-K-3) – fully supported
- Upper-Elementary (CLASS-4 to 6) – in validation
- Secondary (CLASS-S) – in validation
## What Does the CLASS Measure?

<table>
<thead>
<tr>
<th></th>
<th>Emotional Support</th>
<th>Classroom Organization</th>
<th>Instructional Support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-K and K-3</td>
<td>Positive Climate</td>
<td>Behavior Management</td>
<td>Concept Development</td>
</tr>
<tr>
<td></td>
<td>Negative Climate</td>
<td>Productivity</td>
<td>Quality of Feedback</td>
</tr>
<tr>
<td></td>
<td>Teacher Sensitivity</td>
<td></td>
<td>Language Modeling</td>
</tr>
<tr>
<td>Upper Elementary/Secondary</td>
<td>Regard for Student (Adolescent) Perspectives</td>
<td>Instructional Learning Formats</td>
<td>Content Understanding</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Analysis and Problem Solving</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Quality of Feedback</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Instructional Dialogue</td>
</tr>
</tbody>
</table>
Challenging Students to Find the Missing Angle
## CLASS Observations – by domain

### Emotional Support
- Teacher praises students (“Ya’ll did such an excellent job pure thinking”, “I love the way Brandy is tracking”, “You all are the thinkers in the room”) (Positive Climate)
- Teacher uses students’ names (Positive Climate)
- Teacher walks around the classroom and checks in with students (Teacher Sensitivity)
- Teacher encourages students to work in groups (Regard for Student Perspectives)
- Teacher has students share ideas for what the equation is and records these on the board (Regard for Student Perspectives)

### Classroom Organization
- Teacher gives students a time cue (“I’m going to give you two minutes”) (Productivity)
- Teacher sets objective (“You have two minutes to come up with an equation you would write”) (Instructional Learning Formats)
- There is little to no student misbehavior (Behavior Management)

### Instructional Support
- Teacher links to prior learning (“Now that you know the definition from geometry..”) (Content Understanding)
- Teacher gives students the chance to problem-solve by creating an equation to match the missing angle (Analysis and Problem Solving)
- Teacher prompts student to explain how they arrived at an answer (“Can you explain how you got this?”) (Quality of Feedback)
- Teacher prompts students through feedback (“What do you think you should do next?”, “What’s your equation?”, “How can you show that might be right?”) (Quality of Feedback)
- Teacher and student engage in a feedback loop at the end of the clip (Quality of Feedback)
- Students discuss content (Instructional Dialogue)
How is the CLASS organized?

**Instructional Support**

**Content Understanding**

Content Understanding refers to both the depth of lesson content and the approaches used to help students comprehend the framework, key ideas, and procedures in an academic discipline. At a high level, this refers to interactions among the teacher and students that lead to an integrated understanding of facts, skills, concepts, and principles.

<table>
<thead>
<tr>
<th>Low (1,2)</th>
<th>Mid (3,4,5)</th>
<th>High (6,7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth of understanding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Emphasis on meaningful relationships among facts, skills, and concepts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Real world connections</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Multiple and varied perspectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The focus of the class is primarily on presenting discrete pieces of topically related information; broad, organizing ideas are not presented.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The focus of the class is sometimes on meaningful discussion and explanation of broad, organizing ideas, while at other times, it is focused on discrete pieces of topically related information.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The focus of the class is on encouraging deep understanding of content through the provision of meaningful, interactive discussion and explanation of broad, organizing ideas.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Communication of concepts and procedures**

- Essential components identified
- Conditions for how and when to use the concept and/or procedure
- Multiple and varied examples
- Class discussion and materials fail to effectively communicate the essential attributes of concepts/procedures to students.
- Class discussion and materials communicate a few of the essential attributes of concepts/procedures but examples are limited in scope or not consistently provided.
- Class discussion and materials consistently and effectively communicate the essential attributes of concepts/procedures to students.
Standardization
Scoring procedures

• Observe, record notes (15-minute segment)

• Live or video

• Record ratings (1-7) using Manual

• Repeat

• 2-4 cycles in single observation (lesson, 2 hour period)
Training
CLASS Training

- 2 day observation trainings – typically 75-80% of people pass the reliability test; most others pass a second reliability test after follow-up
- Train the Trainer program allows for use at scale
  - Currently over 300 certified trainers
- Over 15,000 people trained to date
  - Currently over 7,000 certified raters
- Online reliability testing and calibration supports – video segments randomized
- Testing online training systems
Reliability
Rater Agreement – 3 studies

- Percent Agreement (within 1)

<table>
<thead>
<tr>
<th></th>
<th>MTP</th>
<th>TUCC - ETS</th>
<th>MET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional Support</td>
<td>86%</td>
<td>89%</td>
<td>77%</td>
</tr>
<tr>
<td>Classroom Organization</td>
<td>85%</td>
<td>86%</td>
<td>83%</td>
</tr>
<tr>
<td>Instructional Support</td>
<td>73%</td>
<td>75%</td>
<td>74%</td>
</tr>
</tbody>
</table>
Validity
Data on CLASS

- CLASS has been used to observe over 20,000 classrooms across the United States
- Great variability in the effectiveness of teachers pre-k to 12
- Teachers with higher scores on CLASS have students who make greater academic and social progress during the school year
- Few students consistently exposed to effective teachers across years (fewer than 10% in elementary school)
Average Ratings of Interactions in Pre-K - 3rd Classrooms

- Emotional Support
- Classroom Organization
- Instructional Support

Class Scores

- Low Quality
- Moderate Quality
- High Quality
CLASS-S Teacher Level Domain Score Distributions

**Emotional Support**

\[ M = 4.03 \]
\[ SD = 0.66 \]

**Classroom Organization**

\[ M = 5.25 \]
\[ SD = 0.62 \]

**Instructional Support**

\[ M = 3.50 \]
\[ SD = 0.60 \]
Average Interactions – PK-8

Pre-K & K (~700 classrooms per grade): NCEDL Study

1st & 3rd (~800 classrooms per grade): NICHD Study of Early Child Care

4th to 8th grade sample (~80 classrooms per grade): Measures of Effective Teaching Study – Gates Foundation

9th Grade – Algebra – ETS Study
Gains in Achievement in Emotionally Supportive Elementary Classrooms

Standardized tests of achievement adjusted

1st Grade Emotional Support

Kindergarten adjustment problems

- No problems
- Multiple problems
CLASS-S results

- Predicts state-standards test scores across all content areas (>2,000 students, >100 classrooms)
  - Moderate effect sizes (.30) in all domains
- Also predict standards tests in the subsequent year
- Now in studies with several thousand classrooms (MET, WTG study).
## Relationship Between CLASS-S and Student Learning

<table>
<thead>
<tr>
<th>Domain</th>
<th>Algebra Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gains</td>
<td></td>
</tr>
<tr>
<td>Emotional Support</td>
<td>.206&lt;sup&gt;t&lt;/sup&gt;</td>
</tr>
<tr>
<td>Classroom Organization</td>
<td>.348&lt;sup&gt;**&lt;/sup&gt;</td>
</tr>
<tr>
<td>Instructional Support</td>
<td>.257&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*Note. ** Significant at the 0.01 level (2-tailed). *Significant at the 0.05 level (2-tailed).*
Alignment with Professional Development
Aligned Professional Development System

- Video library
- MyTeachingPartner
# CLASS Secondary Video Library

## How to use our video library

1. Choose a Dimension
2. Select a video

### Emotional Support
- Positive Climate
- Teacher Sensitivity
- Regard for Adolescent Perspectives

### Classroom Organization
- Behavior Management
- Productivity
- Instructional Learning Formats

### Instructional Support
- Procedures and Skills
- Content Understanding
- Analysis and Problem Solving
- Quality of Feedback

### Student Engagement
- Student Engagement
# Quality of Feedback

<table>
<thead>
<tr>
<th>6 Quality of Feedback Videos Available</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Who Picks the President?</strong></td>
</tr>
<tr>
<td><strong>Working Together to Find an Answer</strong></td>
</tr>
<tr>
<td><strong>Discussing a Poem</strong></td>
</tr>
<tr>
<td><strong>Subtleties in Humor and Seriousness</strong></td>
</tr>
<tr>
<td><strong>Teacher Feedback through Words and Actions</strong></td>
</tr>
<tr>
<td><strong>Capitalizing on an Opportunity for Feedback</strong></td>
</tr>
</tbody>
</table>

**Select another dimension to review:**
- **Positive Climate**
- **Teacher Sensitivity**
- **Regard for Adolescent Persuasive**
- **Behavior Management**
- **Productivity**
- **Instructional Learning Formats**
- **Procedures and Skills**
- **Content Understanding**
- **Analysis and Problem Solving**
- **Quality of Feedback**
- **Student Engagement**

---

**Having trouble viewing videos?**

If you cannot view our CLASS videos, you will need to download the Flash Player. [Download here](#).

If you continue to have problems, please [contact us](#).

---
**Quality of Feedback**

**Discussing a Poem**

**Grade:** Middle School  
**Subject:** History/Social Studies

**Focus Text for this Clip:**
For a lesson about the Harlem Renaissance, this teacher has asked his students to read several poems by Langston Hughes and to identify the one they like the most. In this clip, notice how the teacher continues to focus on the process of learning by asking follow-up questions that require the student to elaborate her thoughts. Watch as the teacher agrees with and extends this student's thoughts by noting that the poem seems to be about larger identity issues, as well as about learning from mistakes. He praises the student in front of her classmates for being able to think about the poem from a broader perspective; providing this sort of specific feedback not only benefits the individual student, but may help other students as well.
Step 1: Teacher videotapes herself interacting with children

Step 2: Coach watches video, writes prompts

Step 3: Teacher reviews prompts & reflects

Step 4: Teacher and coach discuss prompts & practice

Step 5: Coach sends teacher a conference summary and an action plan

The MTP Cycle: 2 weeks
Year 1 Change in Overall Teacher-Student Interactions

**MTPS participation predicts higher quality teacher-student interactions**

![Bar chart showing change in overall teacher-student interactions over pre-test and post-test periods, with MTPS participation predicting higher quality interactions.](chart_image)
Year 1

Intervention Effects on Achievement

- No relation of intervention to either baseline or exit achievement test scores in Year 1 (all $p$’s > .35).

- Why?
  - No evidence we changed the classroom until the very end of the year when most teaching was past.
Year 2
Change in Achievement – Across content areas

MTPS is predicting increases in End of Course Achievement Tests

![Bar chart showing comparison between Pre-Test and Post-Test scores for Control Group.](chart.png)

- Control Group
Pre-K MTP: Improvements in Language Modeling

Language Modeling over time:
- Consultancy: Increasing trend
- Web-Access: Decreasing trend

Dimensions:
- X-axis: September to June
- Y-axis: 3.5 to 5
MTP Effects on Children for Early Career Pre-K Teachers

![Bar chart showing Spring PALS scores over different years of teaching and years of Pre-K education.](chart.png)
Using Observation at Large Scale
Using CLASS at Scale

• **Office of Head Start**
  – monitors all grantees (programs) every 3 years
  – Reauthorization bill required use of a standardized observational measure as part of these monitoring visits
  – Piloted CLASS in 50 grantees in 2008-2009
  – Full implementation in 2009-2010, ongoing
  – OHS provided initial trainings on CLASS for 1 person in every grantee across the country (2400 people), using Train the Trainer model

• **Other Pre-k models:** Georgia Pre-K; California First 5; City of Chicago

• **Quality Rating and Improvement Systems – Statewide –** VA, MN, AZ

• **PK-3 work in Washington, Hawaii, others**

• **PK-12 work in Arlington, VA**
Summary

• CLASS
  – Is an observational tool that produces reliable, valid scores on teacher effectiveness
  – Measures aspects of instruction relevant to student achievement across content areas
  – Can be used in conjunction with content-based measures
  – Can be used at scale
  – Is aligned with a set of professional development materials with demonstrated links to improvement
Acknowledgements

• Colleagues at the **Center for Advanced Study of Teaching and Learning (CASTL)**
  – Robert Pianta
  – Joseph Allen
  – Jason Downer

• Funders:
  – Institute of Education Sciences
  – WT Grant Foundation
  – Bill & Melinda Gates Foundation
More Information

- [http://www.teachstone.com/](http://www.teachstone.com/)
- Bridget K. Hamre
  - hamre@virginia.edu

Thank-you!