



# Using Classroom Observation to Gauge Teacher Effectiveness

## The Classroom Assessment Scoring System (CLASS)

---

Bridget K. Hamre, Ph.D.

Center for Advanced Study of Teaching and Learning  
University of Virginia



# Questions to Answer

---

- Why should we include observations in our assessment of teacher effectiveness?
- What factors should we consider as we design/choose an observational system?
- How can we most effectively and efficiently include observational measures?



# Overview of Today's Presentation

---

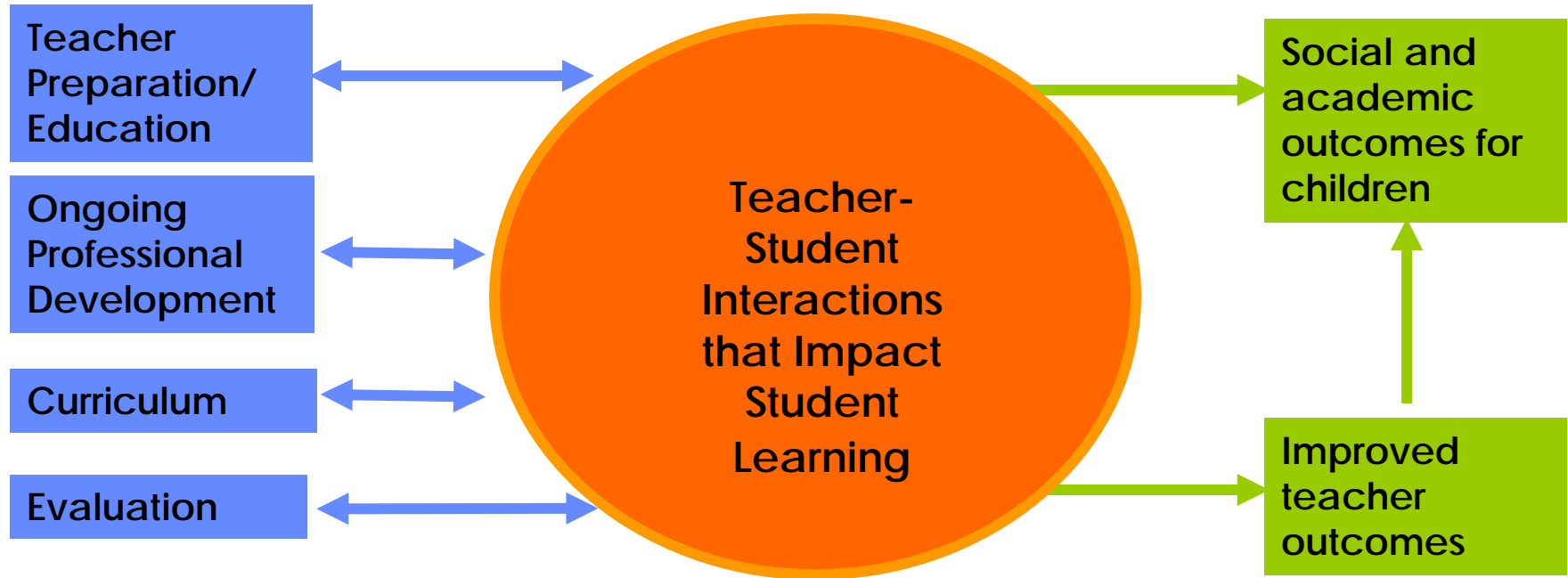
- Classroom observation as one tool for gauging teacher effectiveness
- Factors to consider when choosing an observational tool
- CLASS Overview and Research
- Demonstration of advantages and challenges of observation
- Best practices in conducting observations
- Using observation to improve teaching at scale - examples from the field



***Why include observations?***



# Teacher-Student Interactions as the Active Ingredient in Student Outcomes





# ***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?



# Scope Tradeoffs

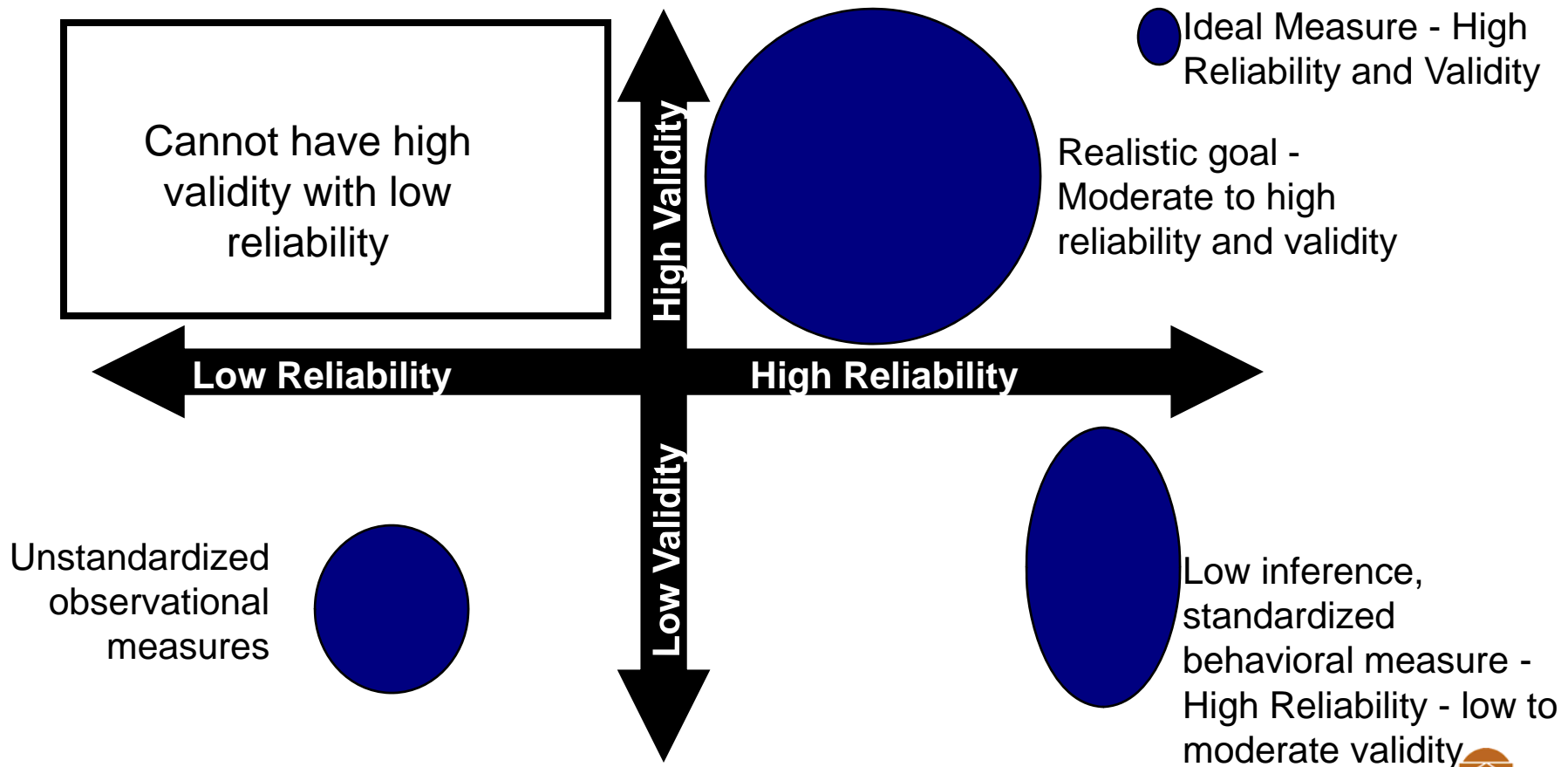
	General	Specific
<b>Grade</b>	PK-12	K, 1, 2, 3, etc
<b>Subject</b>	All	Math only, Algebra only
<b>Student Population</b>	All	ELL; Special Ed
<b>Setting</b>	All	District specific protocol

- General measures offer
  - Ability to use flexibly across schools, classrooms, etc
  - Ability to connect to larger body of data
  - Ability to create shared vision and language
- Specific Measures offer
  - Ability to better assess unique components of settings





# Reliability and Validity Tradeoffs





# 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)?
- **Supplemental Materials:** Does the system include complementary sources of information (such as student surveys, etc.) that could be used to obtain a more complete portrait of the classroom?
- **Links to Improvement:** Does the observation include guidelines and support for using findings for professional development purposes?



# *Overview of CLASS*





# Classroom “Quality”

## STRUCTURE

**What? Who? Where?**

Curriculum

Standards

Materials

Training and Education

## PROCESS

**How?**

Implementation

Relationships

Academic & Social  
Interactions

**Student’s Academic &  
Social Development**



# How Do We Measure Classroom Process?

---

- Classrooms are complex places
- We need a lens through which to view them
- The lens we choose impacts what we see



# The CLASS Lens

- The Classroom Assessment Scoring System (CLASS) offers a lens through which to view classrooms that has several advantages:
  - Based on developmental and learning theories and empirical evidence of classroom interactions that are associated with student outcomes
  - Multiple dimensions of classroom interactions allows for thick description of strengths and weaknesses
  - Evidence that large numbers of people can be trained to use reliably
  - National data on over 3000 classrooms (PK-5) showing that students in classrooms with higher CLASS scores make greater academic and social gains (validity)
  - Intended to be used across grades and content areas
  - Growing body of national data for comparisons
  - Aligned with a professional development system



# Development of CLASS

- NICHD Study of Early Child Care and Youth Development (Classroom Observation System); classroom work started in 1995
- Literature review on empirical articles documenting associations between observable teacher behaviors and student outcomes
- Multiple revisions, pilots, expert reviews, validation studies, etc
- The factor structure was recently validated PK-5 in a sample of 3500 classrooms, nationwide



# CLASS Grade Levels

---

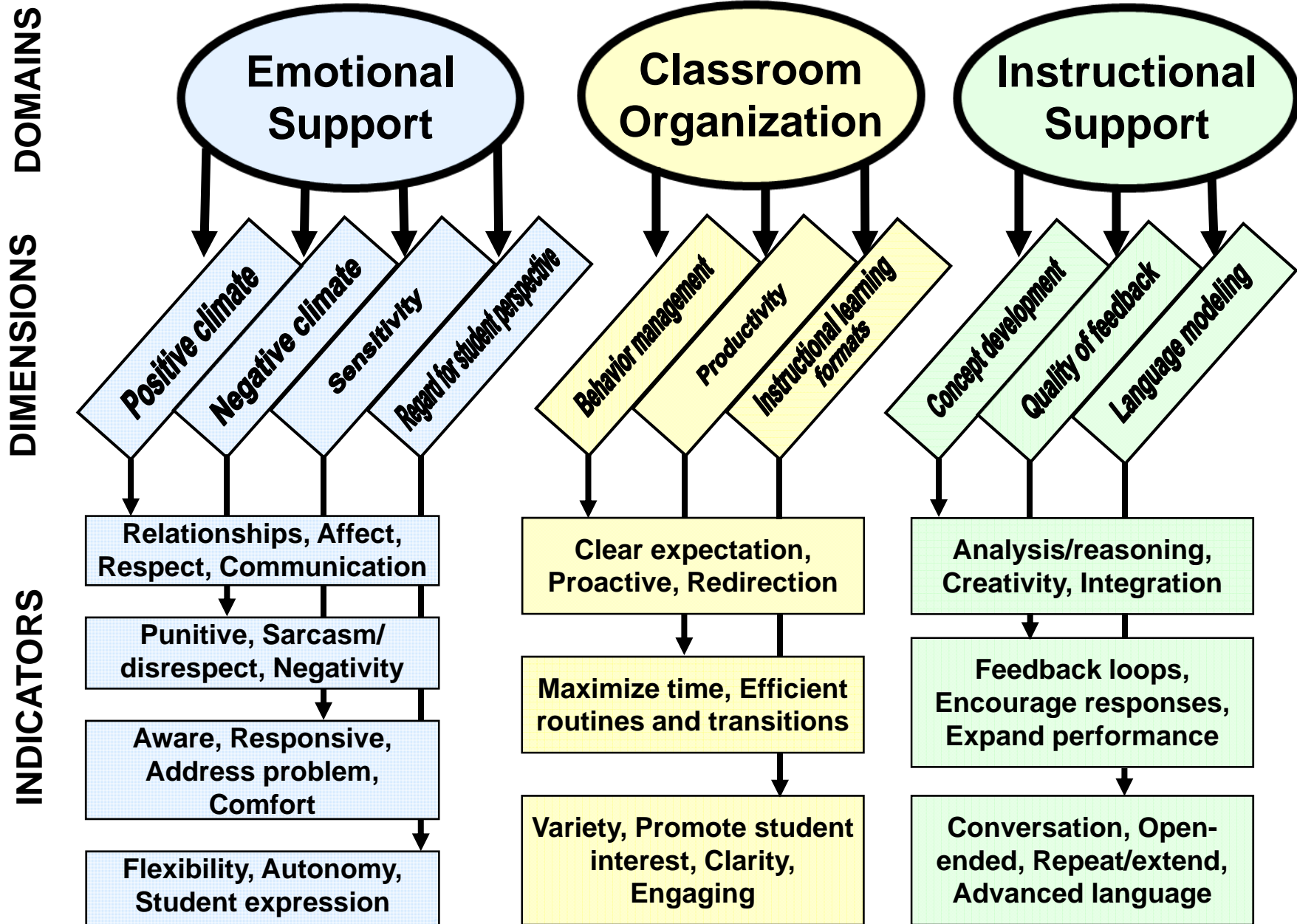
- Preschool: Published 2008
- K-3 (currently used K-5): Published 2008
- Upper Elementary: under development
- Secondary (6-12): Currently in several large validation studies





# What Does the CLASS Measure?

	Emotional Support	Classroom Organization	Instructional Support
<b>Pre-K and Elementary</b>	Positive Climate  Negative Climate	Behavior Management  Productivity	Concept Development Quality of Feedback Language Modeling
<b>Secondary</b>	Teacher Sensitivity  Regard for Student (Adolescent) Perspectives	Instructional Learning Formats	Content Understanding Analysis and Problem Solving Quality of Feedback





# CLASS and Content

---

- CLASS intended for use across content areas
- Assessment of content specific measures suggests more similarity than differences in the types of behaviors that are assessed
- For Pre-k we have developed subject specific measures (Literacy Focus)

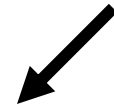


# Scoring - 20 min obs. cycles

1. Take extensive behavioral notes

2. Read your manual

3. Assign a score



<b>Positive Climate</b> Teacher smiles occasionally, kids return smile; teacher asks girl about her weekend; kids somewhat flat at beginning; no evidence of peer connections	1	2	3	4	5	6	7
<b>Negative Climate</b> None observed	1	2	3	4	5	6	7
<b>Teacher Sensitivity</b> Teacher doesn't notice or respond to child with hand raised; T. helps student having hard time tying shoe; S. seem comfortable – raise hands etc;	1	2	3	4	5	6	7
<b>Regard for Student Perspectives</b> Very t. directed lesson; teacher gives frequent reminders for s. to sit on squares; T. not flexible when student asks about going outside; few opportunities for children to lead	1	2	3	4	5	6	7





# CLASS Training

---

- 2 day observation trainings – typically 70-80% 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
- Over 4000 people trained to date



# 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](#)



## CLASS Secondary Video Library

### Quality of Feedback

#### 6 Quality of Feedback Videos Available



[Who Picks the President?](#)



[Working Together to Find an Answer](#)



[Discussing a Poem](#)



[Subtleties in Humor and Seriousness](#)



[Teacher Feedback through Words and Actions](#)



[Capitalizing on an Opportunity for Feedback](#)

Select another dimension to review:

-  [Positive Climate](#)
-  [Teacher Sensitivity](#)
-  [Regard for Adolescent Perspective](#)
-  [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

◀ Go Back to Video List

#### Discussing a Poem

[View Large Size](#)



**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.



# ***CLASS Research - PK-5***



# What observational studies tell us

---

- Exceptional variability within and across grades, generally passive instructional environments
- Little to no association of observed behavior:
  - Teacher experience or training, curriculum, public/private
- Small associations
  - Class size: larger classes more structured; smaller classes more social and higher instructional quality
  - Family income/education related to more positive ratings



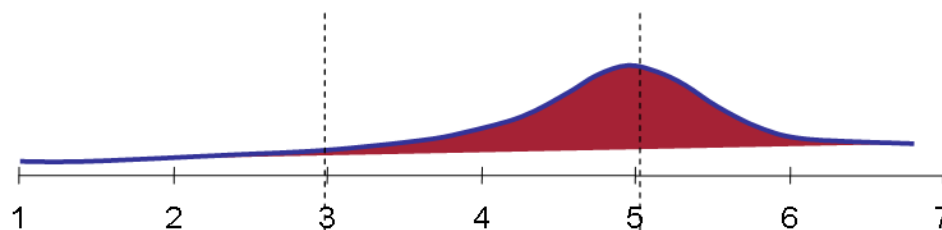
# What observational studies tell us

---

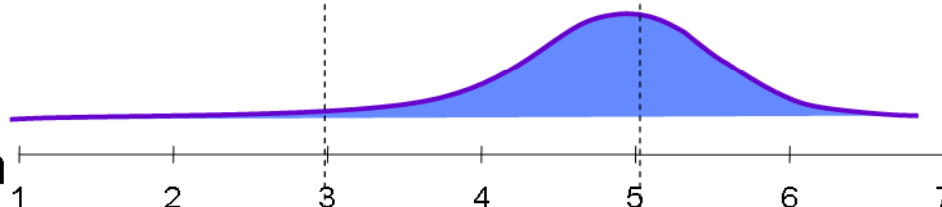
- Students needing access to stable high-quality instruction do not typically receive it – 10% rate
- Students in classrooms with higher emotional and instructional support have display greater academic and social development (particularly for students at-risk of school failure)
- Classroom interactions (as measured by CLASS) are sensitive to intervention – professional development studies (e.g. MyTeachingPartner)

# Average Rating of Interactions in Early Childhood Classrooms (PK-3)

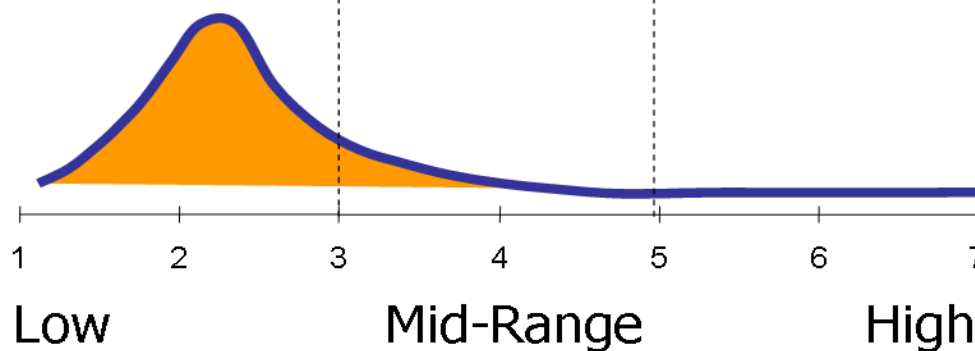
Emotional support



Classroom Organization

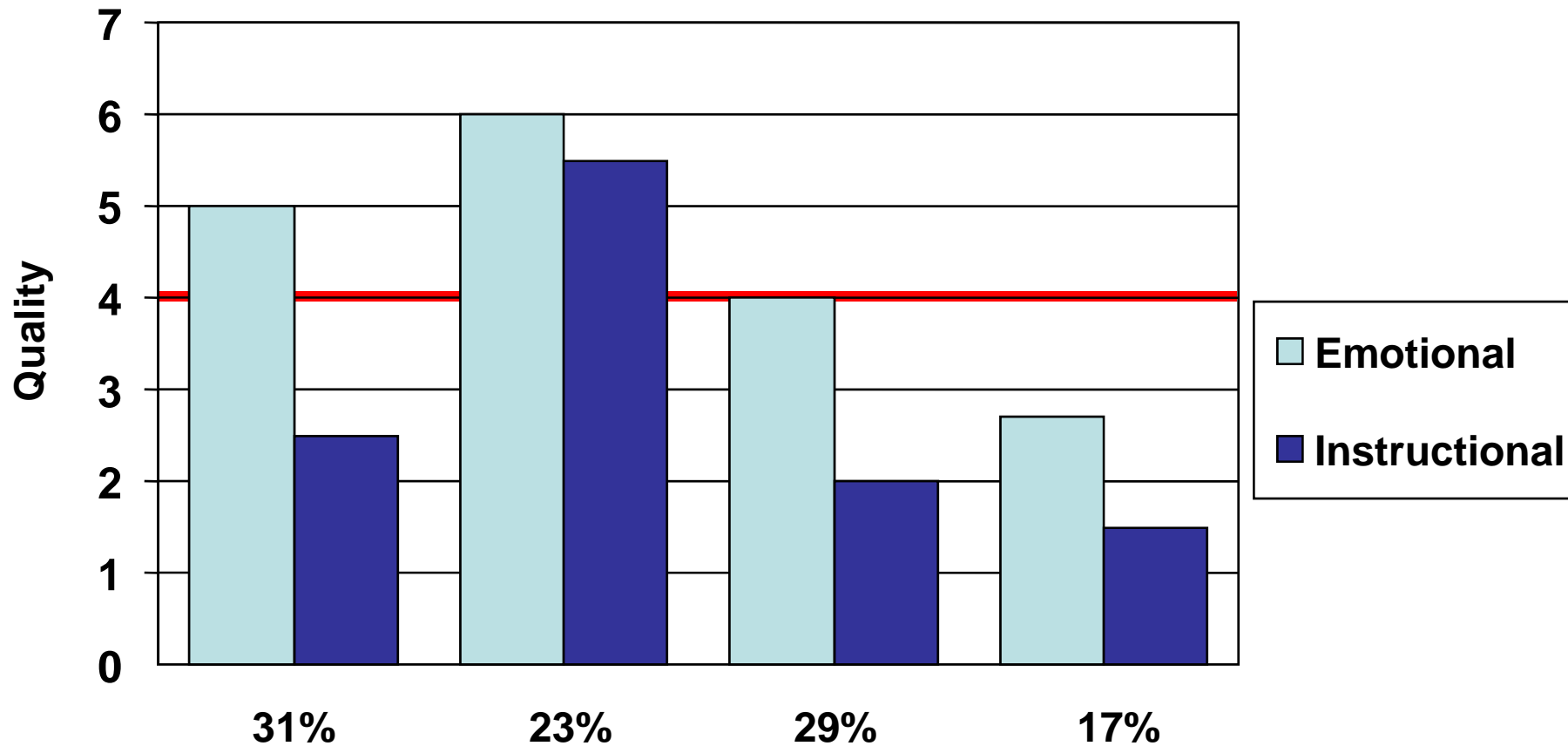


Instructional support

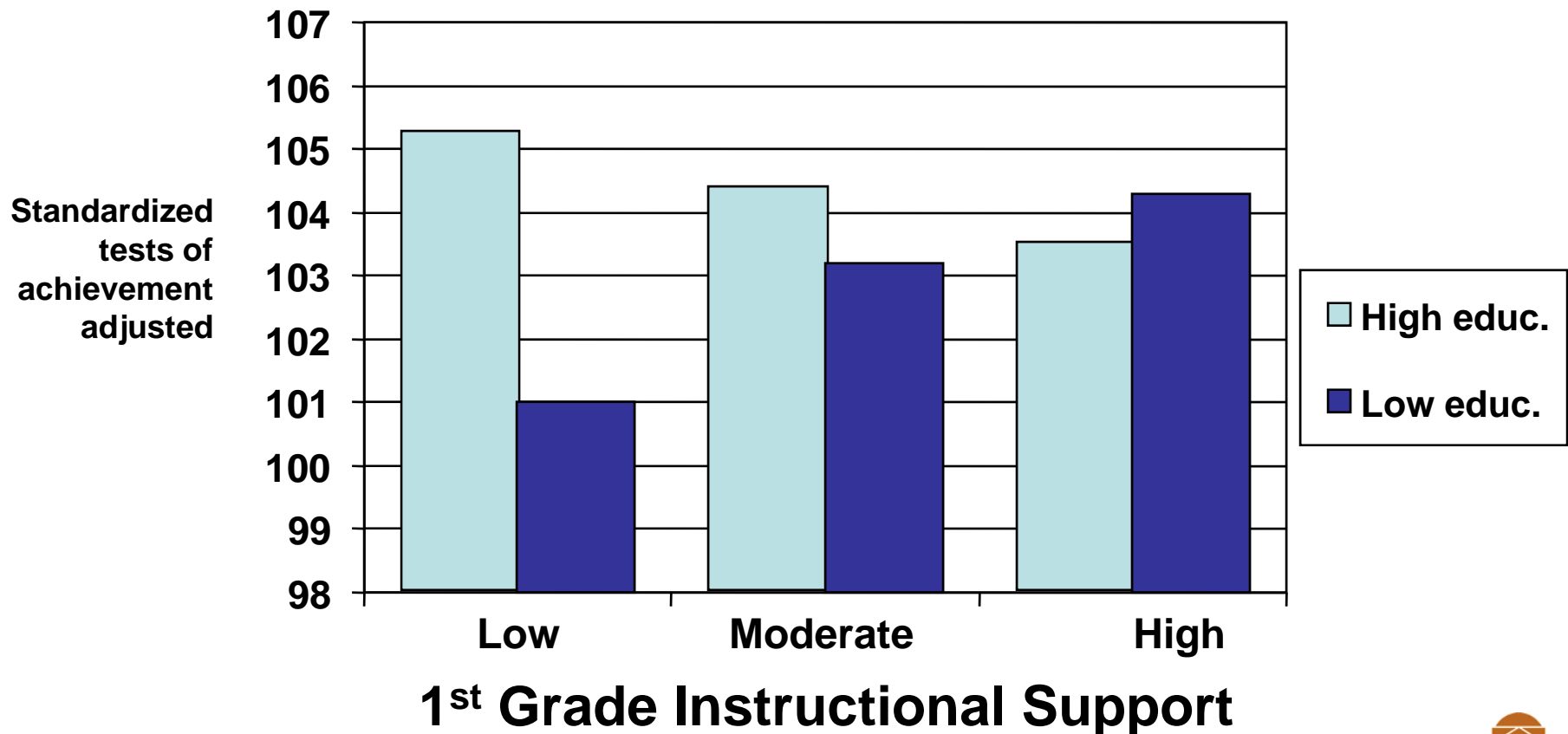




# Profiles of classroom quality: First grade

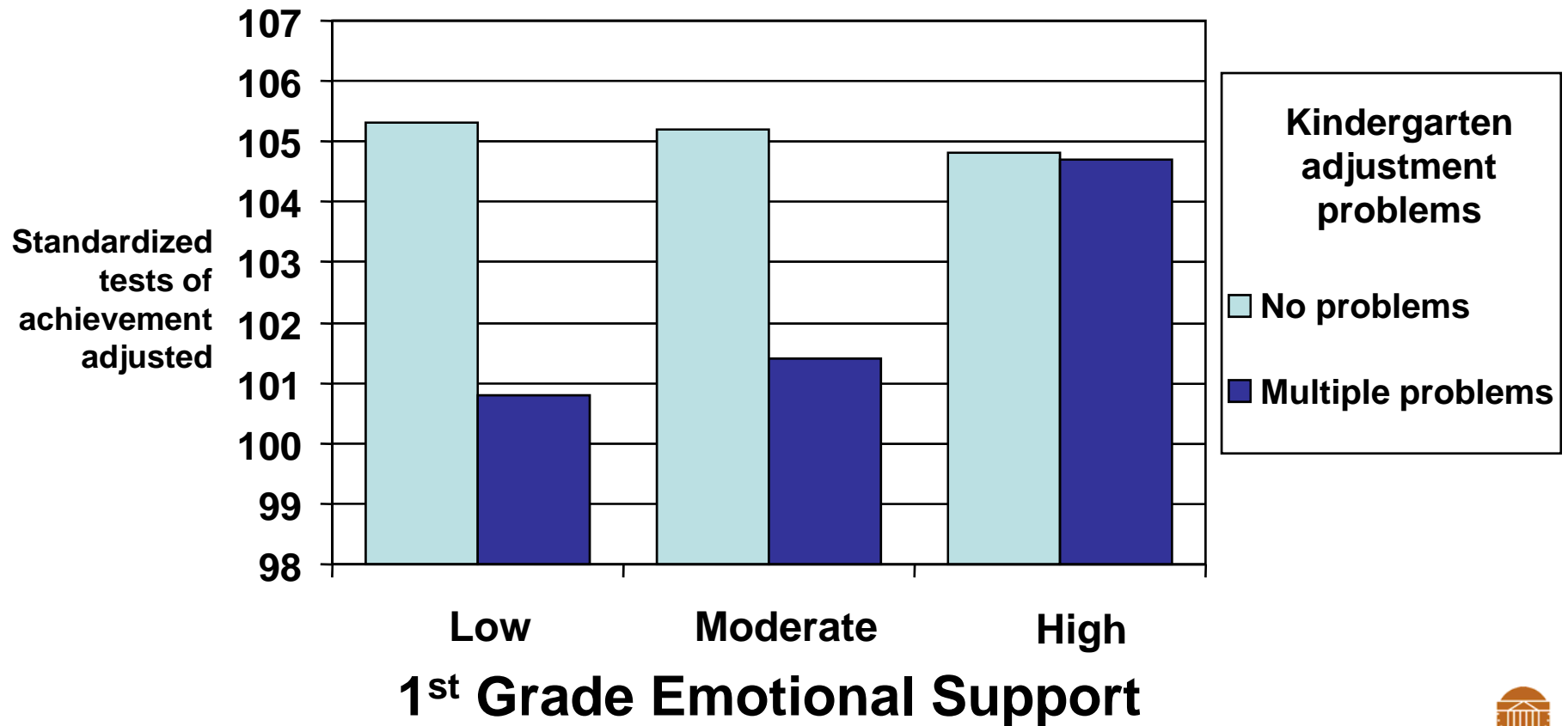


# Gains in grade 1 achievement in instructionally supportive classrooms





# Gains in grade 1 achievement in emotionally supportive classrooms







# New Gates Study

- Attempting to “engineer” a set of tools states and districts can use to assess teacher effectiveness
- Sample in 4 large districts, 250 teachers each in:
  - 3-8 english/language arts and math
  - 9<sup>th</sup> grade algebra
  - 9<sup>th</sup> grade ELA
  - 10<sup>th</sup> grade science (probably biology)
- Will include:
  - Observations (CLASS and several subject specific measures)
  - Value Added
  - Teacher Content Knowledge and Pedagogical Content Knowledge
  - Student assessments of teaching and motivation
- Next year gathering data to examine issues of reliability, validity, feasibility, etc and end up with a system for use at large scale



# ***Demonstration - CLASS Coding***



# Demonstration

---

- Read Regard for Adolescent Perspectives and Quality of Feedback (5 minutes)
- Watch 10 minute video
- Take 10 minutes to give a score for Regard for Adolescent Perspectives and Quality of Feedback
- Take 5 minutes to share scores and talk with group
- Reflections?
  - How much agreement was there among your table?
  - What did you find challenging about using this system?
  - What did you find useful about using this system?
  - Others?



# ***Best Practices***



# Best Practices in Using Observations

- Communication
  - **Recommendations**
    - Make sure people across the system have opportunities to learn about the tool in a meaningful way
    - Be thoughtful about how and by whom results are shared (e.g. do teachers have sufficient knowledge about the tool to interpret results)
    - Do not make inferences that go beyond the scope of the tool
- Issues around Modification
  - **Recommendation**
    - Do not modify standardized tools or you will likely jeopardize reliability and validity



# Timing of Observations (PK-5 only)

---

- Time of day
  - First 30 minutes typically lower, especially on instructional domain
  - Slight decline in Emotional Support across day
- Length of observation
  - 15-20 minute cycles
  - 4 cycles provides fairly reliable estimate (.75-.85 associations with score from 2 full days)
- Time of year
  - Lower in Sept, Dec, and especially May and June
- **Recommendation:** As much as possible observe in consistent ways across classrooms



# Minimizing Rater Effects

- All observational systems will have some rater effects but with planning these can be reduced
- Training
  - Easiest to train on CLASS - grad students, researchers
  - Training should happen close in time to data collection
- Drift/Calibration
  - For data collections lasting more than a few weeks we highly recommend calibration (to measure and each other)
  - Weekly meetings in which everyone codes same tape for check on reliability and feedback
- Multiple raters per classroom or other unit of analysis (more raters better than longer observation time)
- **Recommendation:** Invest in the resources necessary to get good data or you will undermine the possible advantages of including observational measures



# Illustration of Implications of Rater Assignment

Rater A = +1      Rater B = 0      rater C = -1

**Yellow  
School  
5.3**

**Green  
School  
4.0**

**Blue  
School  
6.3**

Class 1	Class 2	Class 3	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
5	4	7	4	4	4	6	6	7

Cheapest option - 1 rater per school	5 (B)	4(B)	7 (B)	5 (A)	5 (A)	5 (A)	5 (C)	5 (C)	6 (C)
	5.3			5			5.3		
Best option - maximum raters per level at of inference	5 (B)	5(A)	6(C)	4 (B)	5 (A)	3 (C)	6 (B)	7 (A)	6 (C)
	5.3			4			6.3		





# Use of Video

---

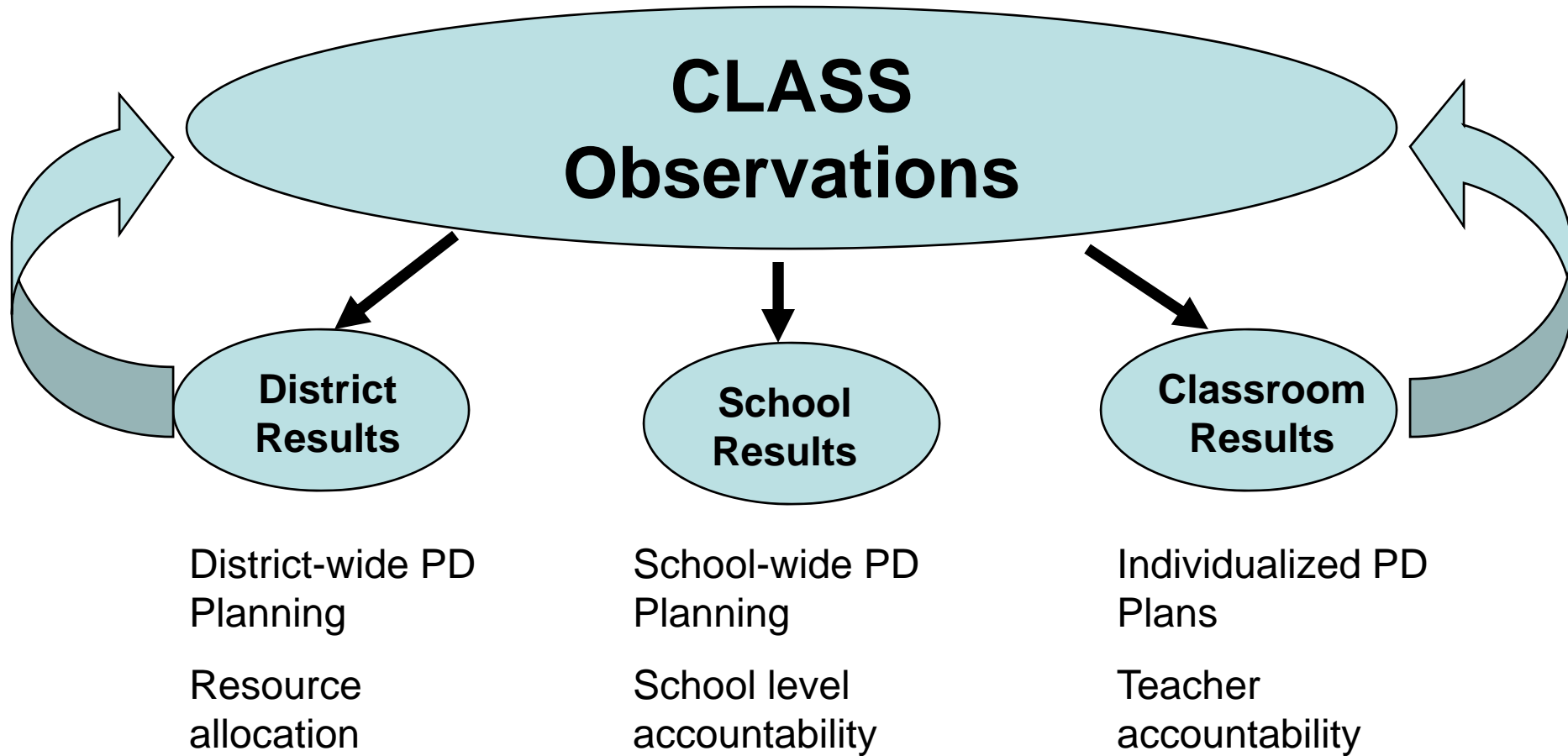
- Advantages:
  - Allows easy way to have multiple raters
  - Allows for documentation of observation for later review and use in professional development efforts
- Disadvantages:
  - Allows for documentation of observation for later review
  - Can be difficult to adequately capture everything that is happening



# ***Using Observation to Promote Change***



# Using Observations to Improve Classroom Practice



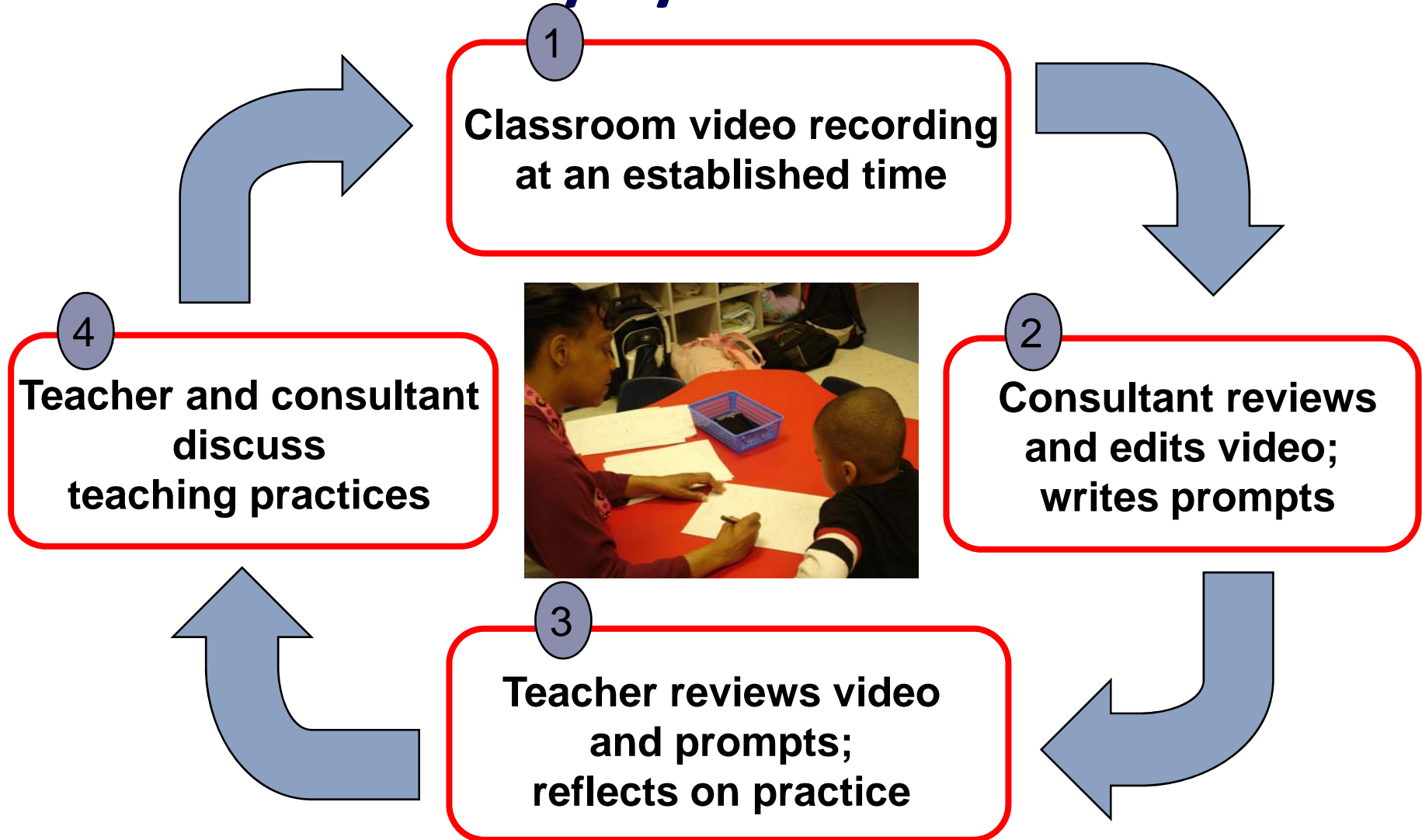


# Examples from the field

---

- Individualized feedback and support – MyTeachingPartner
- Monitoring – Office of Head Start's CLASS pilot - a lever for change

# MTP Consultancy Cycle



## Q

## CLASS Video Library

Welcome to our video library of CLASS constructs. Listed below are the 11 areas of CLASS that we have chosen to focus on and explain further through video demonstrations. You'll get a chance to view teachers interacting with their students in a real-life class setting, while displaying some of the positive behaviors we associate with the various CLASS categories.

Select from the pulldown or pick one of the main categories below.

### CLASS Construct List

**\*\* Use pulldown menu to select**

Positive Climate

GO

#### [Teacher Sensitivity](#)



#### [Positive Climate](#)



#### [Regard for Student Perspectives](#)



#### [Behavior Management](#)



#### [Productivity](#)



#### [Concept Development](#)



#### [Instructional Learning Formats](#)



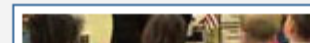
#### [Literacy Focus](#)



#### [Quality of Feedback](#)



#### [Children's Engagement](#)



#### [Language Modeling](#)





# MyTeachingPartner Research

---

- Pre-k teachers were randomly assigned by district to one of two study conditions:
  - **WebAccess** teachers (n=52) received activity descriptions, materials and access to the MyTeachingPartner website
  - **Consultancy** teachers (n=61) received activity descriptions, materials, access to the MyTeachingPartner website, and bi-weekly discussions with a teaching consultant



# Major Findings from 1<sup>st</sup> MTP Study

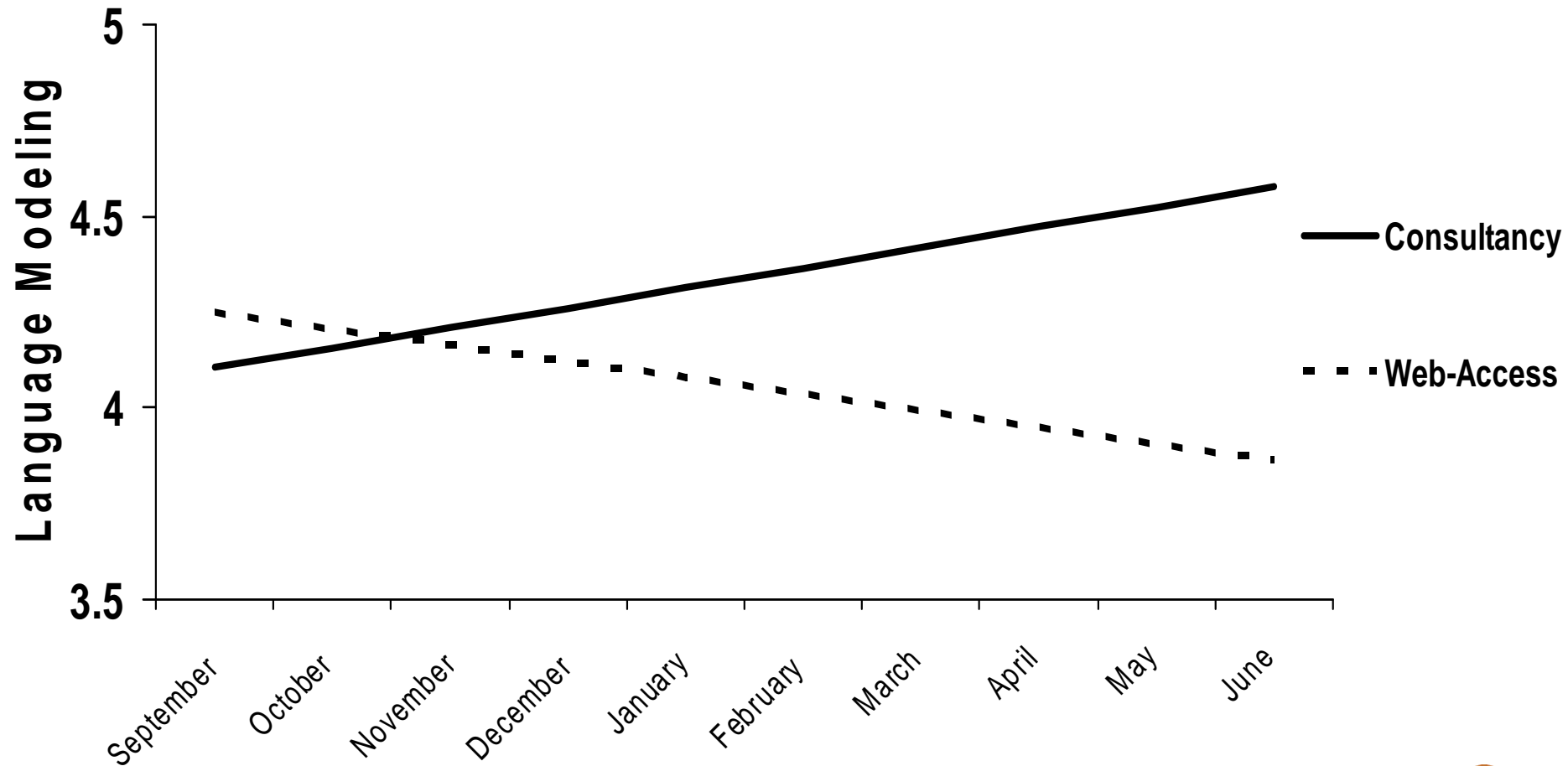
---

- Teachers in consultancy:
  - Grew more sensitive in interactions with students
  - Became more adept at engaging students in instruction
  - Improved the quality of their language stimulation techniques
- Consultancy was a more powerful intervention in high poverty classrooms
- Using web resources was more important in improving quality for teachers without a consultant
- Produced positive gains in student literacy, language and social development, particularly for new teachers



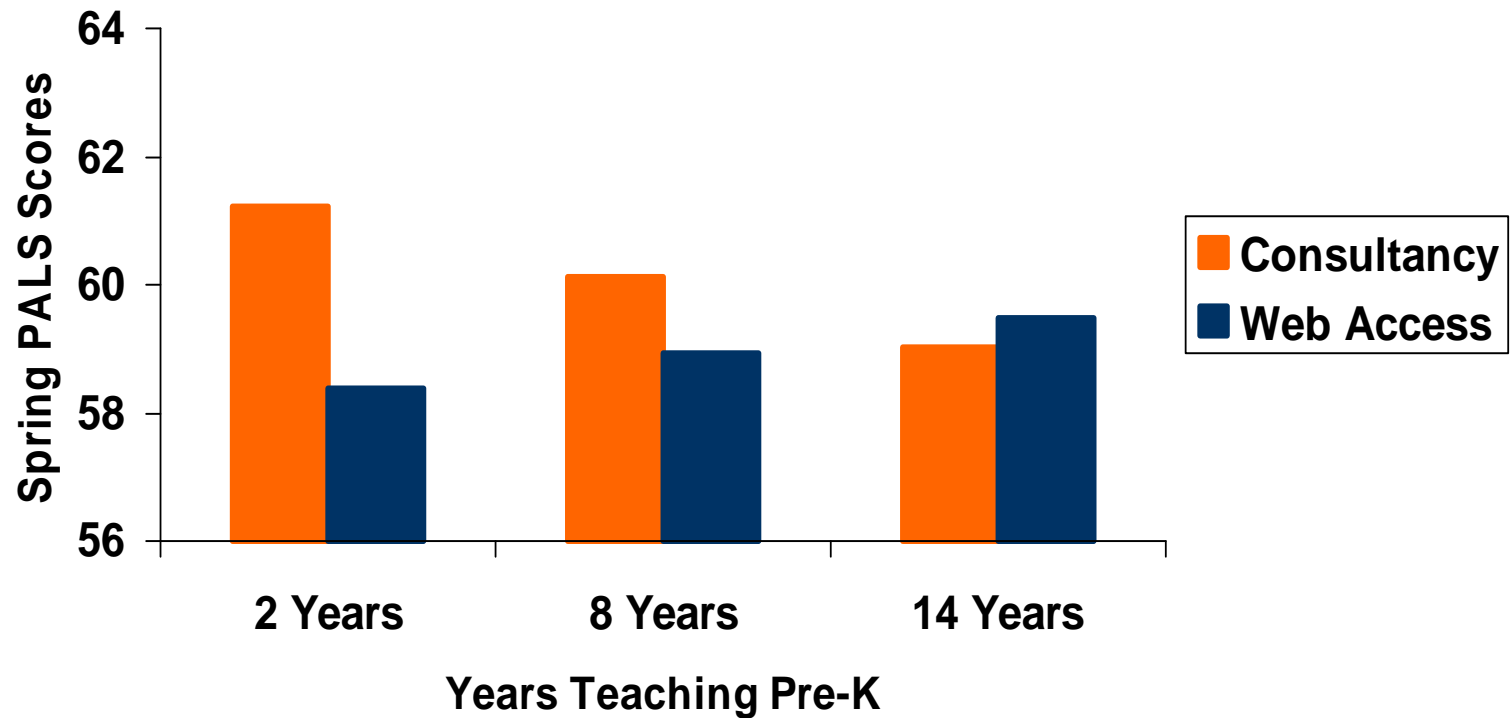


# Improvements in Language Modeling





# Consultation Effects for Early Career Teachers





# Office of Head Start Pilot – 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
- Piloting CLASS in 50 grantees this year
- OHS providing initial trainings on CLASS for 1 person in every grantee across the country (3000 people)



# Questions Answered (in part)

- Why should we include observation in our assessment of teacher effectiveness?
  - Including observational measures provides an important means of assessing teacher effectiveness in way that can promote sustained improvements in practice
  - Use of system-wide observational assessments can help create a shared vision and common vocabulary for discussing effective teaching practices
- What factors should we consider as we design/choose an observational system?
  - Spend time evaluating existing systems based on documented reliability, validity, feasibility, etc.
  - Include all stakeholders in decision making
- How can we most effectively and efficiently include observational measures?
  - Observations, when done well, are not cheap. This is an investment, but one that has the capacity to lead to meaningful organizational change.



# More Information

---

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

# Thank-you!

# Center *for* Advanced Study *of* Teaching *and* Learning

