

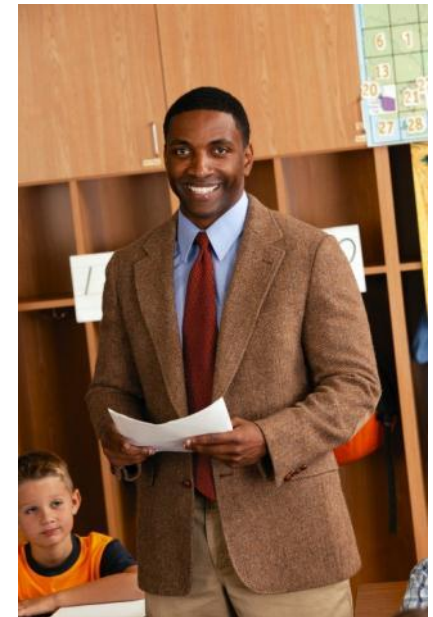
# Welcome

## Northeast Educator Effectiveness Research Alliance (NEERA) Convening

Wednesday, June 5, 2013  
10:30 a.m.–12:00 p.m. ET

Please indicate in the chat:

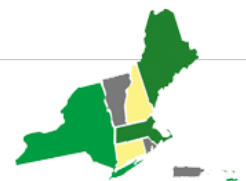
- Your name and affiliation
- The phone number you are calling from



# Agenda



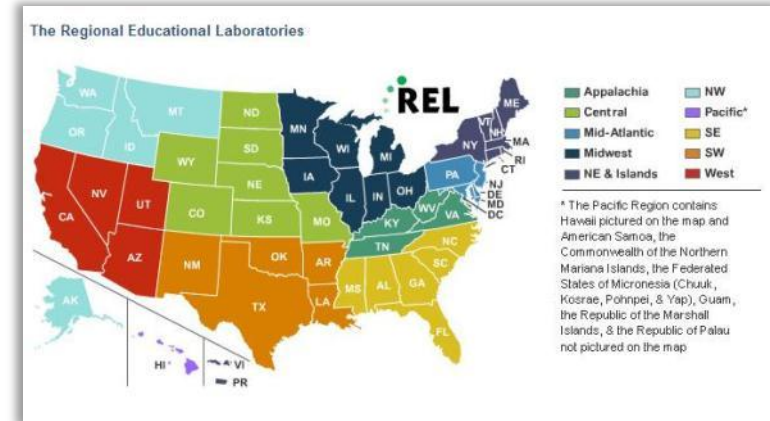
<b>Welcome and Introduction of NEERA</b>	Susan Mundry, Alliance Facilitator, REL-NEI
<b>Evaluator Training Polls</b>	Karen Shakman, Alliance Researcher, REL-NEI
<b>Overview of Evaluator Training: Center on Great Teachers and Leaders</b>	Angela Minnici, Director, Center on Great Teachers and Leaders
<b>Voices from the Field: Evaluator Training in Lynn, MA</b>	Sue Rowe, Lynn Public Schools, MA
<b>Group Discussion</b>	NEERA Membership, facilitated by Susan and Karen
<b>Wrap-Up &amp; Next Steps: NEERA Business</b>	Susan Mundry



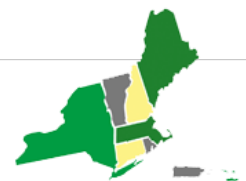
# What Is a Regional Educational Laboratory?



- Network of 10 RELs across the country
- Help states and districts use research and data to inform policy and practice with the goal of improving student outcomes
- Funded by USED Institute of Education Sciences



[ies.ed.gov/ncee/edlabs/](https://ies.ed.gov/ncee/edlabs/)



# Northeast Educator Effectiveness Research Alliance

**Goal:** NEERA provides research to support states' and districts' educator evaluation systems and to build states' and districts' capacity to evaluate their own systems.



Susan Mundry  
*Alliance Facilitator*



Karen Shakman  
*Alliance Researcher*

# NEERA Website



<http://www.relnei.org/research-alliances/northeast-educator-effectiveness-alliance/neera-members.html>

Password: educator



**NORTHEAST & ISLANDS**  
Connecticut • Maine • Massachusetts • New Hampshire • New York  
Puerto Rico • Rhode Island • US Virgin Islands • Vermont



➤ About REL-NEI

➤ About the Region

➤ Regional Priorities

➤ Research Alliances

Technical Assistance

Request Assistance

Works In Progress

Tools & Resources

➤ News

➤ Events

Previous REL Work

Home > Research Alliances > Northeast Educator Effectiveness Research Alliance > NEERA  
Archives and Discussion Forum

## NEERA Archives and Discussion Forum



Welcome to your Research Alliance's password-protected area. Here you may review materials such as meeting minutes and webinar archives, watch for upcoming events, and contribute to our online Discussion Forum with the Alliance team and your fellow members.

[Meeting Materials](#)

[Educator Effectiveness Resources](#)

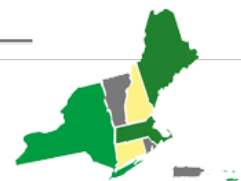
[Upcoming Events](#)

[Discussion Forum](#)

Have any questions or suggestions? Email: [relneiinfo@edc.org](mailto:relneiinfo@edc.org)



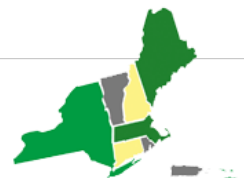
Regional Educational  
Laboratory at EDC  
[relnei.org](http://relnei.org)



# Goals for Today



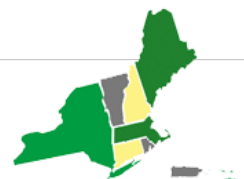
- Learn from experts and practitioners about current practices related to evaluator training/quality assurance
- Learn about needs in the region related to evaluator training/quality assurance
- Continue to recruit new members to NEERA and build cross-state discussions and resource-sharing



# Polling NEERA Members



- Does your state, district, or school require evaluators to participate in a formal training on the new evaluation system?
- Do evaluators in your state, district, or school generally feel prepared to complete all aspects of the evaluation system?
- Check the top three priorities for professional support and development related to evaluator training.

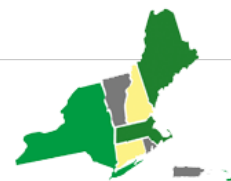


# Early Findings from NH Study



*What are principals saying about the first year of implementing a new teacher evaluation system?*

- Overall positive attitude about goals and intentions of new evaluation system
- Across schools, evaluators' capacity to complete evaluations on schedule varies (#s/evaluator vary)
- Principals/evaluators report need for ongoing training, especially related to SLOs and evaluator calibration.
- Respondents raised some questions about fairness re:
  - consistency of rating (evaluator calibration)
  - rubrics/cut points for particular ratings

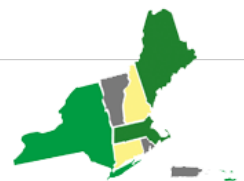




# Voices from the Field



- Sue Rowe  
Lynn, MA Public Schools



# Evaluator Training: Ensuring Quality and Reliability

Angela Minnici

Director, Center on Great Teachers & Leaders  
American Institutes of Research

Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

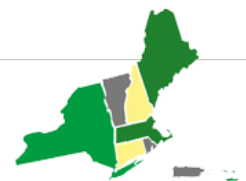


# Agenda



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Overview of state approaches to training
- Challenges and solutions to:
  - Rubric design
  - Personnel
  - Interrater reliability



# Adaptive or Technical Challenge?



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■



“Indeed, the single most common source of leadership failure we’ve been able to identify...is that people, especially those in positions of authority, treat adaptive challenges like technical problems.”

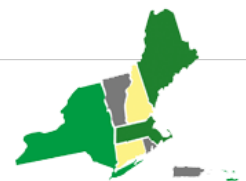
—Heifetz & Linsky (2002), p. 14.

# Technical Versus Adaptive Challenges



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Technical Challenges
  - Can be fixed by experts and by implementation of best practices.
  - Are easy to identify and have solutions that can be implemented quickly.
- Adaptive Challenges
  - Require people to change their values, behaviors, and attitudes.
  - Require people with the problem to do the work of solving it.
  - Can take longer to implement.





Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

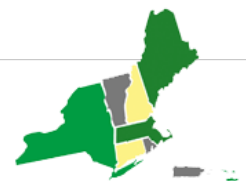
# State Approaches



Regional Educational  
Laboratory at EDC  
[relnei.org](http://relnei.org)



Major Partners:  
 **AIR** WestEd 



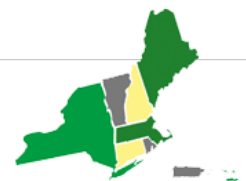
# Types of Training



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

Training approaches for educator evaluation vary both in **format** (i.e., how it's delivered) and **content** (i.e., what is provided).

- **District leadership training**
  - Statewide conference
  - Online videos or webinars
- **Educator orientation training**
  - Train-the-trainer
  - Online self-paced webinars
- **Evaluator training**
  - In person
  - Online
  - Blended





Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

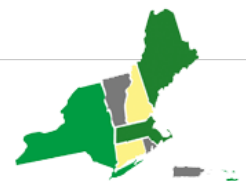
# Rubric Design



Regional Educational  
Laboratory at EDC  
[relnei.org](http://relnei.org)



Major Partners:  
 **AIR**  **WestEd**





# A Well Designed Rubric Can Improve...



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

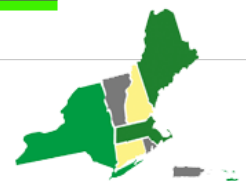
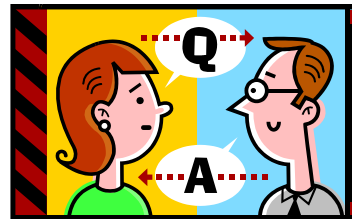
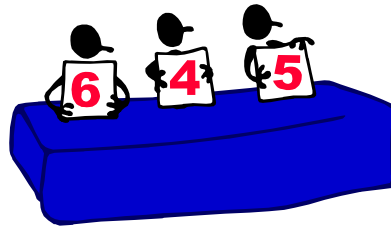
- Scoring



- Feedback



- Professional growth



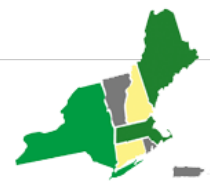
# Types of Rubrics



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Observation Rubric
  - Evidence from single or multiple observations
- Professional Practice Rubric
  - Observation evidence
  - Artifact Review
  - Other Measures (e.g. surveys)
- Evaluation Rubric
  - Professional Practice evidence
  - Student growth evidence
  - Goal-setting and/or professional development (if applicable)

*Note: Rubrics may also include scoring metrics or notes*



# Ensuring High Quality Design



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

## Common problems:

- A lack of vertical alignment
- Vague language
- A lack of distinctness between domains and indicators

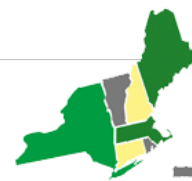
## Can be addressed by:

- Stakeholder engagement
- High quality training
- Clear descriptions and examples
- Alignment to leading standards or other rubrics (InTASC, Danielson)
- Expert Review

Component 2a: Creating an Environment of Respect and Rapport			
LEVEL		CRITICAL ATTRIBUTES	POSSIBLE EXAMPLES
4	Classroom interactions among the teacher and individual students are highly respectful, reflecting genuine warmth and caring and sensitivity to students as individuals. Students exhibit respect for the teacher and contribute to high levels of civility among all members of the class. The net result of interactions is that of connections with students as individuals.	In addition to the characteristics of a level of performance 3, <ul style="list-style-type: none"> <li>Teacher demonstrates knowledge and caring about individual students' lives beyond school.</li> <li>When necessary, students correct one another in their conduct towards classmates.</li> <li>There is no disrespectful behavior among students.</li> <li>The teacher's response to a student's incorrect response respects the student's dignity.</li> </ul>	<ul style="list-style-type: none"> <li>Teacher inquires about a student's soccer game last weekend</li> <li>Students say "Shhh" to classmates while the teacher or another student is speaking.</li> <li>Students clap enthusiastically for one another's presentations for a job well done.</li> <li>The teacher says: "That's an interesting idea, Student J, but you're 'forgetting....'"</li> <li>And others...</li> </ul>
3	Teacher-student interactions are friendly and demonstrate general caring and respect. Such interactions are appropriate to the ages, of the students. Students exhibit respect for the teacher. Interactions among students are generally polite and respectful. Teacher responds successfully to disrespectful behavior among students. The net result of the interactions is polite and respectful, but impersonal.	<ul style="list-style-type: none"> <li>Talk between teacher and students and among students is uniformly respectful.</li> <li>Teacher responds to disrespectful behavior among students.</li> <li>Teacher makes superficial connections with individual students.</li> </ul>	<ul style="list-style-type: none"> <li>Teacher greets students by name as they enter the class or during the lesson.</li> <li>The teacher gets on the same level with students, such as kneeling beside a student working at a desk.</li> <li>Students attend fully to what the teacher is saying.</li> <li>Students wait for classmates to finish speaking before beginning to talk.</li> <li>Students applaud politely following a classmate's presentation to the class.</li> <li>Students help each other and accept help from each other.</li> <li>Teacher and students use courtesies such as please/thank you, excuse me.</li> <li>Teacher says: "Don't talk that way to your classmates" and the insults stop.</li> <li>And others...</li> </ul>
2	Patterns of classroom interactions, both between the teacher and students and among students, are generally appropriate but may reflect occasional inconsistencies, favoritism, and disregard for students' ages, cultures, and developmental levels. Students rarely demonstrate disrespect for one another. Teacher attempts to respond to disrespectful behavior, with uneven results. The net result of the interactions is neutral: conveying neither warmth nor conflict.	<ul style="list-style-type: none"> <li>The quality of interactions between teacher and students, or among students, is uneven, with occasional disrespect.</li> <li>Teacher attempts to respond to disrespectful behavior among students, with uneven results.</li> <li>Teacher attempts to make connections with individual students, but student reactions indicate that the efforts are not completely successful or are unusual.</li> </ul>	<ul style="list-style-type: none"> <li>Students attend passively to the teacher, but tend talk, pass notes, etc. when other students are talking.</li> <li>A few students do not engage with others in the classroom, even when put together in small groups.</li> <li>Students applaud half-heartedly following a classmate's presentation to the class.</li> <li>Teacher says: "Don't talk that way to your classmates" but student shrugs his/her shoulders</li> </ul>
1	Patterns of classroom interactions, both between the teacher and students and among students, are mostly negative, inappropriate, or insensitive to students' ages, cultural backgrounds, and developmental levels. Interactions are characterized by sarcasm, put-downs, or conflict. Teacher does not respond to disrespectful behavior.	<ul style="list-style-type: none"> <li>Teacher uses disrespectful talk towards students; Student body language indicates feelings of hurt or insecurity.</li> <li>Students use disrespectful talk towards one another with no response from the teacher.</li> <li>Teacher displays no familiarity with or caring about individual students' interests or personalities.</li> </ul>	<ul style="list-style-type: none"> <li>A student slumps in his/her chair following a comment by the teacher.</li> <li>Students roll their eyes at a classmate's idea; the teacher does not respond.</li> <li>Many students talk when the teacher and other students are talking; the teacher does not correct them.</li> <li>Some students refuse to work with other students.</li> <li>Teacher does not call students by their names. I agree about the sequence; let's do this on the next pass.</li> </ul>

Framework for Teaching - Copyright 2011, Outcomes Association, Inc. All rights reserved. May not be incorporated into an electronic platform.

3



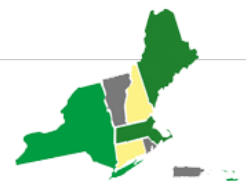
# Examples



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- [Rhode Island Professional Practice Rubric](#)
  - Description of overall performance, indicators, and examples
  - Online platform for documentation
- [Ohio Teacher Performance Evaluation Rubric](#)
  - Evidence sources specified
  - Evidence documentation
- [Colorado Rubric for Evaluation Teachers](#)
  - Checklist across performance levels
  - Differentiation between observable and unobservable indicators

***All have descriptions of performance at each level***





Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

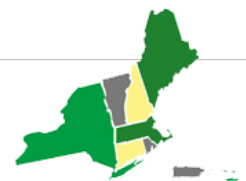
# Personnel



Regional Educational  
Laboratory at EDC  
[relnei.org](http://relnei.org)



Major Partners:  
 **AIR**  **WestEd**

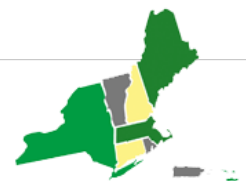


# The Role of the Evaluator



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

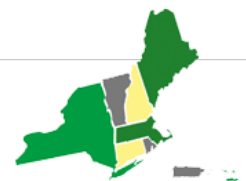
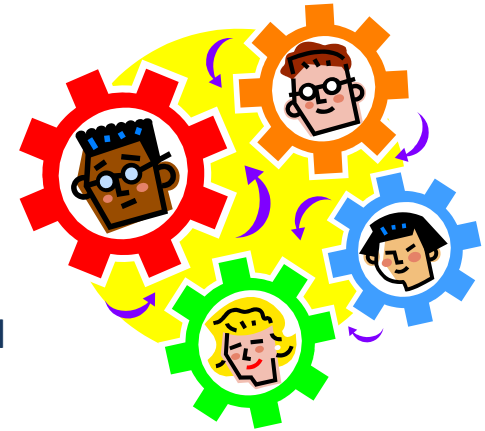
- At the heart of an evaluation system are personnel who can accurately assess teacher performance, communicate the results of that assessment to teachers, and help them plan for their professional growth.
- *Who the observers are is less important than whether they*
  - *Receive adequate training*
  - *Receive ongoing feedback on their performance*
  - *Periodically reassess and calibrate their observation skills*



# The Role of the Evaluator



- Challenges for principals as evaluators:
  - Limited time
  - Limited background in certain subjects.
  - Lack of teaching experience
  - Limited skills and experience with conducting professional conversations and coaching
- Although the principal should conduct the final summative evaluation, other trained staff may observe teachers:
  - Principal
  - Vice-principal
  - Content-area administrators
  - Teacher leaders
  - Peer Observers (full-time or part-time)

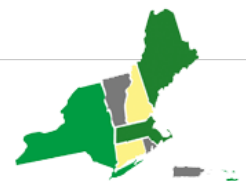


# Training Personnel to be Evaluators



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- High-quality training builds trust.
- High-quality training needs to include:
  - Certification exams and calibration exercises
  - How to provide high-quality feedback
- Evaluation and feedback are a professional responsibility for principals.
  - What responsibilities can other staff assist with (i.e. observation, feedback)?
  - How do teachers and principals connect evaluation results to professional development and learning?







Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

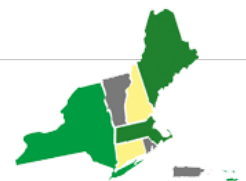
# Interrater Reliability



Regional Educational  
Laboratory at EDC  
[relnei.org](http://relnei.org)



Major Partners:  
 **AIR**  **WestEd**

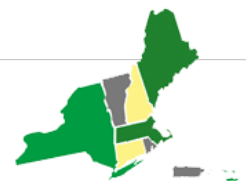


# Assessing Validity in Teacher Evaluation



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Examine the relationship between teacher practice and student learning
  - Teacher practice measure: classroom observation ratings
  - Student learning measure: teacher-level added value (student growth)
- A valid teacher observation instrument
  - Low observation ratings correlate with low value-added scores
  - High observation ratings correlate with high value-added scores



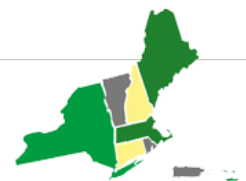
# Importance of Interrater Reliability



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

Even with a terrific observation instrument, the results are meaningless if observers are not trained to agree on evidence and scoring.

→ A teacher should get the same score no matter who observes him or her.

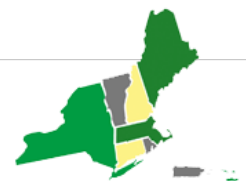


# Interrater Reliability



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Interrater reliability is one element of an observational system:
  - Instruments
  - Raters
  - Scoring designs
- Variability may influence teacher scores:
  - Teachers
  - Lessons
  - Raters
- There is not a single right metric for interrater agreement.
- Generalizability studies can help can assist in the design of cost-efficient systems that produce reliable scores (Hill et al., 2012).

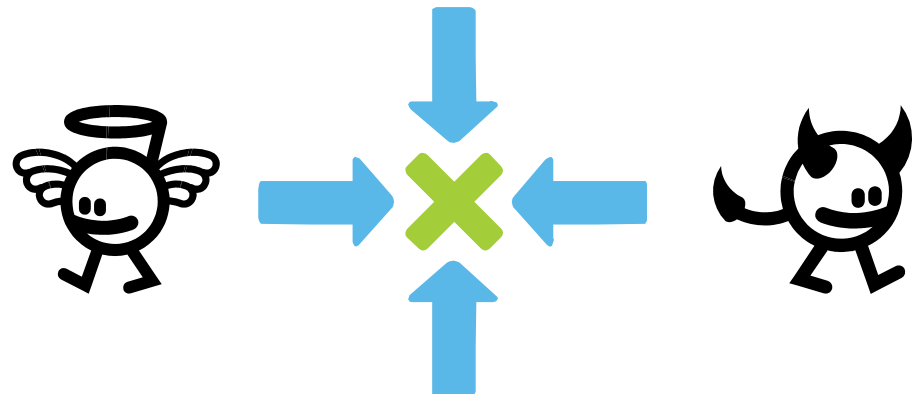


# Obstacles to Rater Accuracy



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

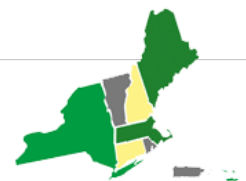
- Rater bias
- Leniency or Severity
- Central tendency
- Halo or Horns effects



High-quality training

Formal calibration and certification

Informal calibration through meetings



# The Importance of Multiple Observers and Observations



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Using multiple observers—and multiple observations—improves the reliability of scores more than having longer observation periods.

(Sources: Ho & Kane 2013; Sartain et al., 2011)



# Reliability Results With Various Combinations of Raters and Number of Lessons



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

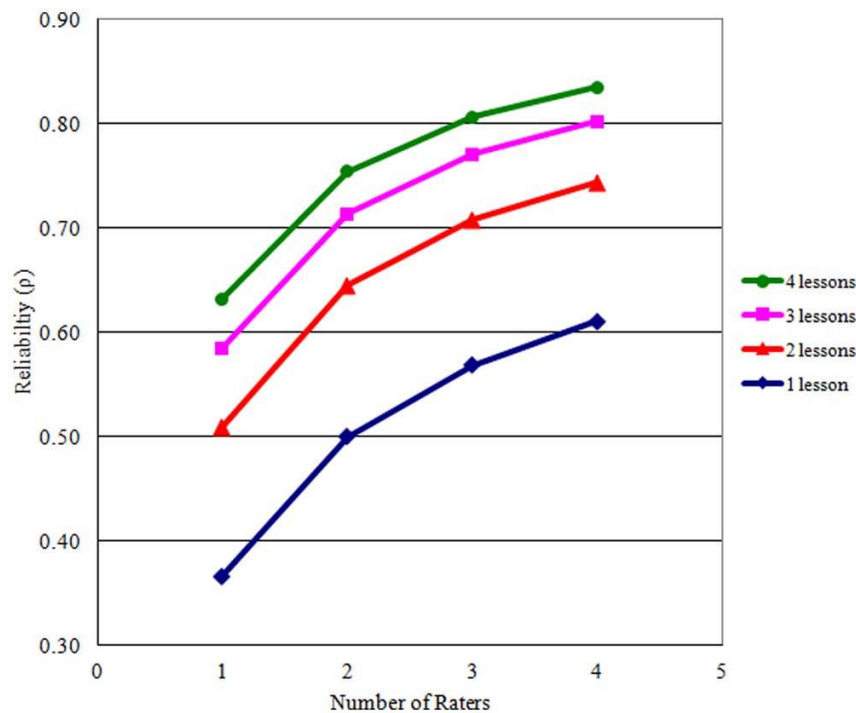
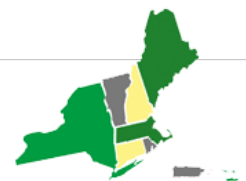


Figure 2. *Errors and Imprecision: The Reliability of Different Combinations of Raters and Lessons.* From Hill et al. (2012). Used with permission of author.

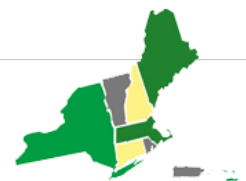


# Lessons From the Measures of Effective Teaching (MET) Study



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Interrater reliability depends on factors beyond teacher quality, such as the consistency of classroom context, student demographics, and differences between lessons.
- Rater severity isn't an issue for the majority of observers but should be monitored
- The authors of the study recommend that
  - Observers undergo training and calibration prior to scoring.
  - Teachers be observed multiple times.
  - The district employ impartial observers from outside the school.



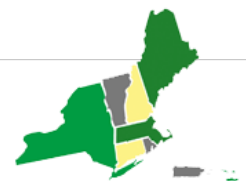


# Lessons From the Measures of Effective Teaching (MET) Study



Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■

- Evaluation outcomes were most valid when they combined
  - Student feedback (surveys)
  - Student learning (growth and/or achievement)
  - Observation
- The most valid way of combining these measures was to weight them comparably as part of a teacher's overall evaluation.

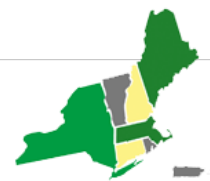


# References



- Heifetz, R. A., & Linsky, M. (2002). *Leadership on the line: Staying alive through the dangers of leading*. Cambridge, MA: Harvard Business School Press.
- Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher*, 41(2), 56–64.  
<http://scholar.harvard.edu/mkraft/publications/when-rater-reliability-not-enough-teacher-observation-systems-and-case-g-study>
- Ho, A. D., & Kane, T. J. (2013). The reliability of classroom observations by school personnel. Seattle, WA: Bill and Melinda Gates Foundation.  
[http://www.metproject.org/downloads/MET\\_Reliability%20of%20Classroom%20Observations\\_Research%20Paper.pdf](http://www.metproject.org/downloads/MET_Reliability%20of%20Classroom%20Observations_Research%20Paper.pdf)

Center on  
**GREAT TEACHERS & LEADERS**  
at American Institutes for Research ■



Center on Great Teachers and Leaders  
1000 Thomas Jefferson Street NW  
Washington, DC 20007-3835  
877-322-8700  
[www.gtlcenter.org](http://www.gtlcenter.org)  
[gtlcenter@air.org](mailto:gtlcenter@air.org)

▶ *Advancing state efforts to grow, respect, and retain great teachers  
and leaders for all students*

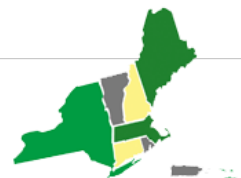
# Discussion: Breakout Groups



# NEERA Business



- Update on NH Study: data collection nearly completed
- Upcoming NEERA workshops 2013
  - Practitioner Data Use: June 10 and June 26, 2:00–4:30
  - CPG Spring Phone Conference: June 26, 10:30-noon
  - SLO Bridge Event: September 17 or 18
  - Program and Policy Evaluation Workshop: TBD, Fall 2013



# Contact Us



- Alliance Facilitator: Susan Mundry—  
[smundry@wested.org](mailto:smundry@wested.org)
- Alliance Researcher: Karen Shakman —  
[kshakman@edc.org](mailto:kshakman@edc.org)

Thank You!