

# Evaluating Teacher Effectiveness: Some Models to Consider

**Laura Goe, Ph.D.**

**Ohio Task Force**

**Columbus, Ohio ♦ January 26, 2011**

Copyright © 2009 National Comprehensive Center for Teacher Quality. All rights reserved.

# The “other 69%”

- In one state, 69% of teachers (Prince et al., 2006) could not be accurately assessed with VAMs
  - Teachers in subject areas that are not tested with annual standardized tests
  - Teachers in grade levels (lower elementary) where no prior test scores are available
  - Questions about the validity of measuring special education teachers and ELL teachers with VAMs
- May be higher in many states—75-80%

# A few teacher evaluation models

- TAP (Teacher Advancement Program)
- Austin, TX
- Rhode Island
- Washington, DC
- Delaware

# Questions to ask about models

- Do they include measures that are “rigorous and comparable across classrooms”?
- Do they include measures that show learning growth “between two points in time”?
- Do they include measures that are aligned with and focused on grade level and subject standards?
- Do they allow teachers from all subjects to be evaluated with evidence of student learning growth?
- Will using this model help improve teaching and learning?

# Teacher Advancement Program (TAP) Model

- TAP requires that teachers in tested subjects be evaluated with value-added models
- All teachers are observed in their classrooms (using a Charlotte Danielson type instrument) at least three times per year by different observers (usually one administrator and two teachers who have been appointed to the role)
- Teacher effectiveness (for performance awards) determined by combination of value-added and observations
- Teachers in non-tested subjects are given the school-wide average for their value-added component, which is combined with their observation scores

# Austin Independent School District

## Student Learning Objectives:

- Teachers determine two SLOs for the semester/year
- One SLO must address all students, other may be targeted
- Use broad array of assessments
- Assess student needs more directly
- Align classroom, campus, and district expectations
- Aligned to state standards/campus improvement plans
- Based on multiple sources of student data
- Assessed with pre and post assessment
- Targets of student growth
- Peer collaboration

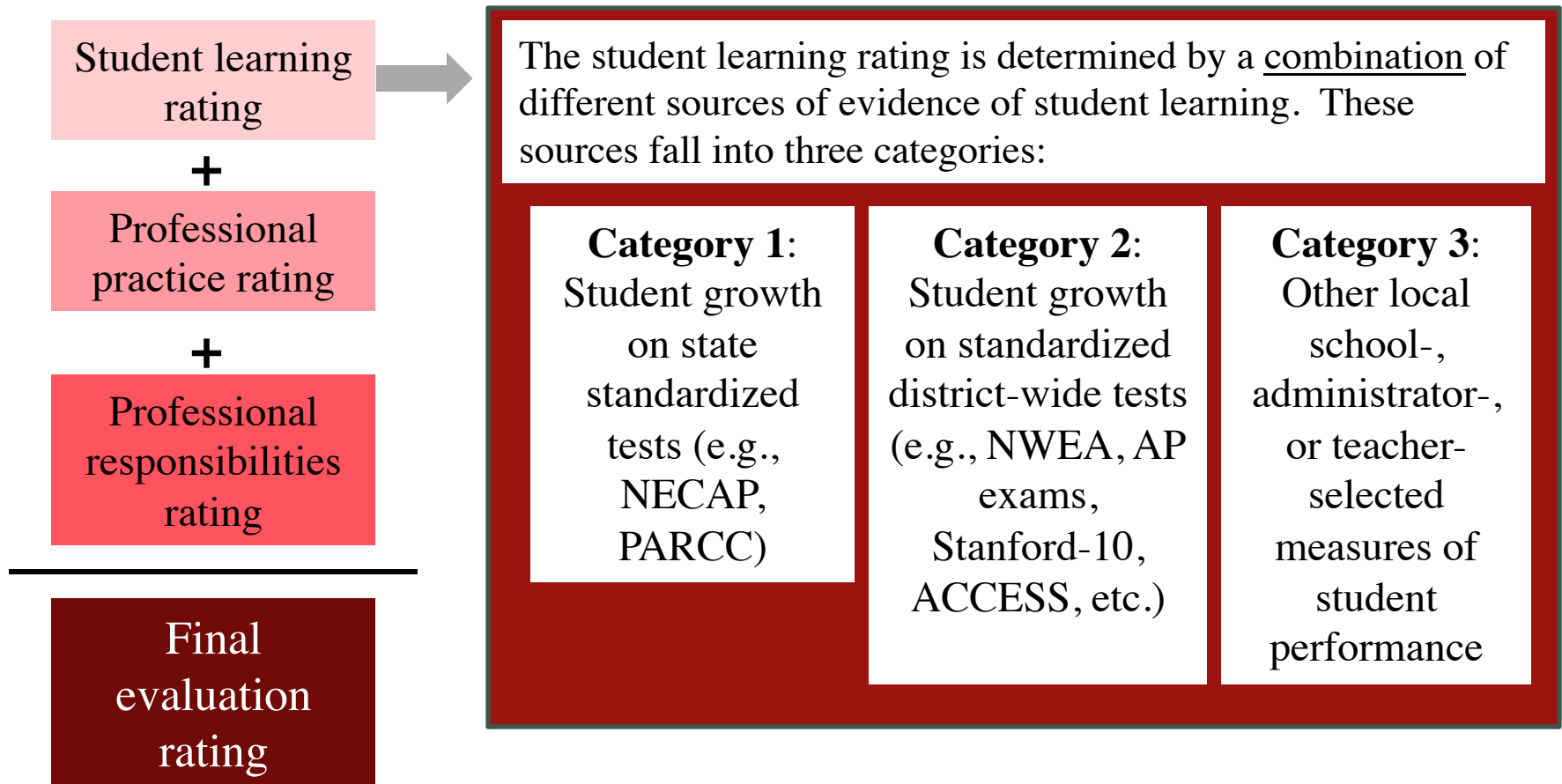
# Rubric for student learning objectives

<i>APPROVAL</i>		<i>NEEDS REVISION</i>	
<i>LEVEL 4</i>	<i>LEVEL 3</i>	<i>LEVEL 2</i>	<i>LEVEL 1</i>
<i>Indicates expert use of data, rigorous goal-setting based on student and community strengths and needs, incorporates appropriate valid/reliable assessments, considers teacher's own prior performance, and demonstrates alignment with campus improvement goals.</i>	<i>Indicates good use of data and acceptable level of rigor for students, considers teacher's prior performance, and utilizes appropriate assessments, but is not aligned with broader campus goals and does not incorporate the strengths of students or the school community.</i>	<i>Suggests superficial use of data, is minimally rigorous for students and the teacher, uses related but unproven assessments, and does not incorporate the goals of the campus or the strengths of students or the school community.</i>	<i>Represents ineffective use of data, is not rigorous for students or the teacher, does not support the campus goals, utilizes inappropriate assessment, and does not consider strengths of students or the school community.</i>
<b>Needs Assessment/Rationale:</b> What are the needs?			
<ul style="list-style-type: none"> <li>■ Campus data are reviewed for areas of strength and need (within subject area, within grade level, within student group, examining the Campus Improvement Plan (CIP), etc.).</li> <li>■ Classroom data are reviewed for areas of strength and need (by subject area, by student group, by concepts/skills/behavior).</li> </ul>	<ul style="list-style-type: none"> <li>■ Classroom data are reviewed for areas of strength and need (by subject area, by student group, by concepts/skills/behavior).</li> <li>■ Campus needs and strengths are not incorporated.</li> </ul>	<ul style="list-style-type: none"> <li>■ Classroom data are reviewed for areas of need, but needs of specific student groups are not examined in depth, and strengths are not identified.</li> <li>■ Campus needs and strengths are not incorporated.</li> </ul>	<ul style="list-style-type: none"> <li>■ Classroom data are not used to identify student needs.</li> <li>■ Campus needs and strengths are not incorporated.</li> </ul>
<b>Learning Content/Context and Student Group:</b> What and who is targeted?			
<ul style="list-style-type: none"> <li>■ Targets specific academic concepts, skills, or behaviors based on TEKS/TAKS Objective</li> <li>■ Targets the needs of the identified population</li> <li>■ Considers demonstrated strengths of identified population, as well as classroom &amp; school community</li> <li>■ Targets year-long (or semester-long) concepts, skills, or behaviors</li> <li>■ Supports goals of the Campus Improvement Plan (CIP)</li> </ul>	<ul style="list-style-type: none"> <li>■ Targets specific academic concepts, skills, or behaviors based on TEKS/ TAKS Objective</li> <li>■ Targets the needs of the identified population.</li> <li>■ Considers demonstrated strengths of identified population, as well as classroom &amp; school community.</li> <li>■ Targets year-long (or semester-long) concepts, skills, or behaviors.</li> <li>■ CIP Goals are not incorporated.</li> </ul>	<ul style="list-style-type: none"> <li>■ Targets specific academic concepts, skills, or behaviors based on TEKS/TAKS Objective</li> <li>■ Targets year-long (or semester-long) concepts, skills, or behaviors.</li> <li>■ Does not target the needs of all students in the identified population.</li> <li>■ Does not consider the strengths of the identified population, classroom, or school community.</li> <li>■ CIP Goals are not incorporated.</li> </ul>	<ul style="list-style-type: none"> <li>■ Does not target concepts, skills, or behaviors based on TEKS/TAKS Objective</li> <li>■ Does not target year-long (or semester-long) concepts, skills, or behaviors.</li> <li>■ Does not target the needs of all students in the identified population.</li> <li>■ Does not consider the strengths of the identified population, classroom, or school community.</li> <li>■ CIP Goals are not incorporated.</li> </ul>

# Rubric for student learning objectives (cont'd)

<i>Level 4</i>	<i>Level 3</i>	<i>Level 2</i>	<i>Level 1</i>
<b>Learning Objective:</b> What will students learn?			
<ul style="list-style-type: none"> <li>■ Based on the identified student needs.</li> <li>■ Supports goals of the CIP.</li> <li>■ Is rigorous.</li> <li>■ Is a good example of ongoing, reflexive practice.</li> <li>■ Provides clear focus for instruction and assessment.</li> <li>■ Is measurable.</li> <li>■ Reflects strengths of students and school community.</li> </ul>	<ul style="list-style-type: none"> <li>■ Based on the identified student needs.</li> <li>■ Is rigorous.</li> <li>■ Is a good example of ongoing, reflexive practice.</li> <li>■ Provides clear focus for instruction and assessment.</li> <li>■ Is measurable.</li> <li>■ Reflects strengths of students and school community.</li> <li>■ Does not relate to goals of CIP.</li> </ul>	<ul style="list-style-type: none"> <li>■ Based on the identified student needs of some students.</li> <li>■ Provides clear focus for instruction and assessment.</li> <li>■ Is measurable.</li> <li>■ Does not reflect strengths of students and school community.</li> <li>■ Does not relate to goals of CIP.</li> <li>■ Is not rigorous.</li> <li>■ Is not a good example of ongoing, reflexive practice.</li> </ul>	<ul style="list-style-type: none"> <li>■ Does not consider student needs.</li> <li>■ Does not provide a clear focus for instruction.</li> <li>■ Is not measurable.</li> <li>■ Does not reflect strengths of students and school community.</li> <li>■ Does not relate to goals of CIP.</li> <li>■ Is not rigorous.</li> <li>■ Is not a good example of ongoing, reflexive practice.</li> </ul>
<b>Outcome Assessment:</b> How will you know whether they learned it?			
<ul style="list-style-type: none"> <li>■ Aligns with the targeted learning content area.</li> <li>■ Relationship with learning objective is apparent.</li> <li>■ Has been demonstrated as reliable and valid for targeted students.</li> <li>■ Follows guidelines for appropriate assessments.</li> </ul>	<ul style="list-style-type: none"> <li>■ Aligns with the targeted learning content area.</li> <li>■ Relationship with learning objective is apparent.</li> <li>■ Follows guidelines for appropriate assessments.</li> <li>■ Has not been demonstrated as reliable and valid for targeted students.</li> </ul>	<ul style="list-style-type: none"> <li>■ Aligns with the targeted learning content area.</li> <li>■ Relationship with learning objective is minimally apparent.</li> <li>■ Does not follow guidelines for appropriate assessments.</li> <li>■ Has not been demonstrated as reliable and valid for targeted students.</li> </ul>	<ul style="list-style-type: none"> <li>■ Does not align with the targeted learning content area.</li> <li>■ Does not follow guidelines for appropriate assessments.</li> <li>■ Has not been demonstrated as reliable and valid for targeted students.</li> </ul>
<b>Performance Target:</b> What is your goal for student achievement?			
<ul style="list-style-type: none"> <li>■ Predicts performance based on past performance of students when available.</li> <li>■ Is a rigorous expectation for students.</li> <li>■ Is a rigorous expectation for teachers, based on past performance.</li> </ul>	<ul style="list-style-type: none"> <li>■ Predicts performance based on past performance of students when available.</li> <li>■ Is a rigorous expectation for students.</li> <li>■ Does not “stretch” the teacher beyond past performance history.</li> </ul>	<ul style="list-style-type: none"> <li>■ Predicts performance based on past performance of students when available.</li> <li>■ Is not a rigorous expectation for students.</li> <li>■ Does not “stretch” the teacher beyond past performance history.</li> </ul>	<ul style="list-style-type: none"> <li>■ Does not utilize past performance of students when available.</li> <li>■ Is not a rigorous expectation for students.</li> <li>■ Does not “stretch” the teacher beyond past performance history.</li> </ul>
<b>Rigor:</b> How rigorous is your SLO?			
<ul style="list-style-type: none"> <li>■ Content is challenging, complex and progressively deepens knowledge of core content</li> <li>■ Content is thought-provoking requiring high thinking demand</li> <li>■ Requires analytical thinking and active use of knowledge</li> <li>■ Content is relevant to life/ experiences.</li> </ul>	<ul style="list-style-type: none"> <li>■ Content is challenging and complex for most, but not all, students.</li> <li>■ Requires analytical thinking.</li> <li>■ Content is relevant for most, but not all, students.</li> </ul>	<ul style="list-style-type: none"> <li>■ Content is challenging for some students.</li> <li>■ Does not require analytical thinking.</li> <li>■ Content is relevant for some students.</li> </ul>	<ul style="list-style-type: none"> <li>■ Content is not challenging.</li> <li>■ Does not require analytical thinking.</li> <li>■ Is not relevant to life and learning experiences.</li> </ul>

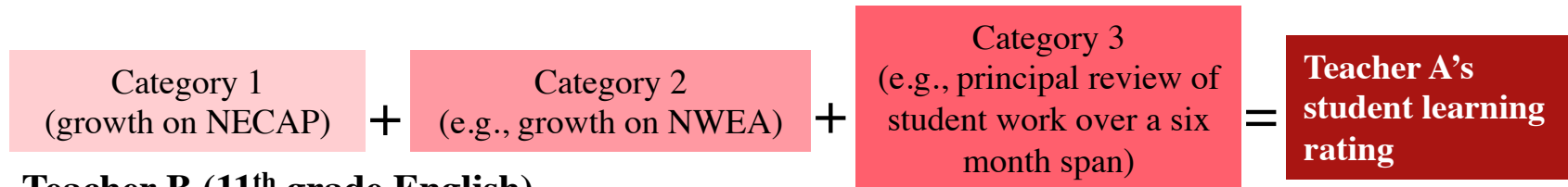
# Rhode Island DOE Model: Framework for Applying Multiple Measures of Student Learning



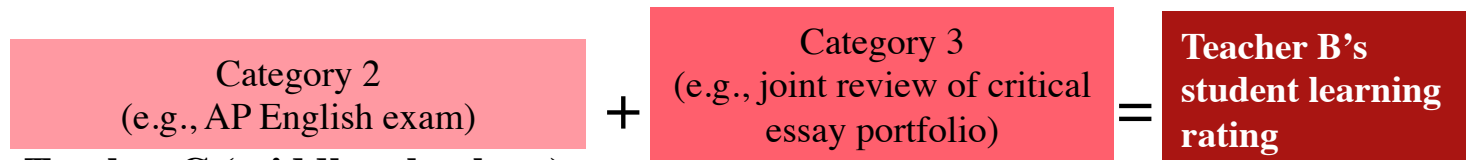
# Rhode Island Model: Student Learning Group Guiding Principles

- “Not all teachers’ impact on student learning will be measured by the same mix of assessments, and the mix of assessments used for any given teacher group may vary from year to year.”

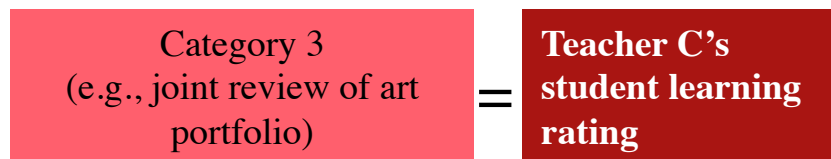
## Teacher A (5<sup>th</sup> grade English)



## Teacher B (11<sup>th</sup> grade English)



## Teacher C (middle school art)



*This teacher may use several category 3 assessments*

# Washington DC's IMPACT: Teacher Groups

- Group 1: general ed teachers for whom value-added data can be generated
- Group 2: general ed teachers for whom value-added data *cannot* be generated
- Group 3: special education teachers
- Group 4: non-itinerant English Language Learner (ELL) teachers and bilingual teachers
- Group 5: itinerant ELL teachers
- Etc...

# Score comparison for Groups 1 & 2

	Group 1 (tested subjects)	Group 2 (non-tested subjects)
Teacher value-added (based on test scores)	50%	0%
Teacher-assessed student achievement (based on non-VAM assessments)	0%	10%
Teacher and Learning Framework (observations)	35%	75%
Commitment to School Community	10%	10%
School Wide Value-Added	5%	5%

# Non-VAM tests (accepted under Washington, DC's IMPACT evaluation system)

- DC Benchmark Assessment System (DC BAS)
- Dynamic Indicators of Basic Early Literacy Skills (DIBELS)
- Developmental Reading Assessment (DRA)
- Curriculum-based assessments (e.g., Everyday Mathematics)
- Unit tests from DCPS-approved textbooks
- Off-the-shelf standardized assessments that are aligned to the DCPS Content Standards
- Rigorous teacher-created assessments that are aligned to the DCPS Content Standards
- Rigorous portfolios of student work that are aligned to the DCPS Content Standards

# Hillsborough, FL

- Stated goal is to evaluate **every** teacher's effectiveness with student achievement growth, even teachers in non-tested subjects and grades
- Undertaking to create pre- and post-assessments for all subjects and grades
- Expanding state standardized tests and using value-added to evaluate more teachers
- Part of a multiple measures system

# Delaware Model

- Standardized test will be used as part of teachers' scores in some grades/subjects
- "Group alike" teachers, meeting with facilitators, determine which assessments, rubrics, processes can be used in their subjects/grades (multiple measures)
- Assessments must focus on standards, be given in a "standardized" way, i.e., giving pre-test on same day, for same length of time, with same preparation
- Teachers recommend assessments to the state for approval
- Teachers/groups of teachers take primary responsibility for determining student growth
- State will monitor how assessments are "working"

# Thoughts on identifying measures to be used in non-tested subjects

- Teachers are the subject/grade experts
- They are already measuring student learning progress for many purposes
- Have them meet in subject/grade alike teams to identify measures that would be appropriate to assess students progress toward the appropriate standards
- A panel could discuss the recommendations and approve measures that meet criteria

- Evaluating the models group activity
- Report out on activity

# The “caseload” teachers

- For nurses, counselors, librarians, and other professionals who do not have their own classroom, what counts is their “caseload”
  - May be all the students in the school
  - May be a specific set of students
  - May be other teachers
  - May be all of the above!
  - Other outcomes may need to be considered for them (not just student achievement growth)

# Models

Austin (Student Learning Objectives)

<http://www.austinisd.org/inside/initiatives/compensation/slos.phtml>

Teacher Advancement Program

<http://www.tapsystem.org/>

Washington DC IMPACT Guidebooks

[http://www.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+\(Performance+Assessment\)/IMPACT+Guidebooks](http://www.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)/IMPACT+Guidebooks)

Rhode Island Model

[http://www.ride.ri.gov/educatorquality/EducatorEvaluation/Docs/Working%20Group%](http://www.ride.ri.gov/educatorquality/EducatorEvaluation/Docs/Working%20Group%20)

Delaware Model

[http://www.doe.k12.de.us/csa/dpasii/student\\_growth/default.shtml](http://www.doe.k12.de.us/csa/dpasii/student_growth/default.shtml)

# Questions?





**Laura Goe, Ph.D.**

**P:** 609-734-1076

**E-Mail:** lgoe@ets.org

**National Comprehensive Center for  
Teacher Quality**

1100 17th Street NW, Suite 500

Washington, DC 20036-4632

877-322-8700 > [www.tqsource.org](http://www.tqsource.org)