

Race to the Top Teacher and Leader Effectiveness/Standards and Assessment Community of Practice Convening

A Meeting of the SLO Work Group

April 25, 2012

Boston, MA

How We Got Here

- Survey administered to States in March
- Feedback reveals strong State interest in **1) developing systems to control for the quality of SLOs and assessments and 2) co-creating/refining tools/documents to support SLOs**
- Current plan based on State feedback
- Meeting agenda and objectives for today are based on State feedback

Remember Your Dropbox Tool!

The RSN has already delivered to you the most complete library of SLO tools/documents sorted by category ever assembled

Meeting Goal

To lay the groundwork for a successful SLO work group that addresses state needs and results in tangible products

Meeting Objectives

- To build understanding across the work group of where states stand with SLO implementation, what their assets are and the challenges they face
- To review the draft work group plan and prioritize deliverables
- To advance state understanding of SLO quality control and quality assessment procurement and development mechanisms
- To identify specific tools/documents states would like to co-create/refine over the next 90-120 days
- To identify states to participate in work group subgroups that design a quality control framework for SLO implementation and a protocol to help evaluators calibrate SLOs for quality

Agenda

8:00 am – 8:35 am	Welcome, Self-Assessment, Introductions
8:35 am – 9:05 am	Review SLO Work Group Plan, Prioritize Deliverables
9:05 am – 9:35 am	Presentation on Quality Control – Leigh McGuigan
9:35 am – 10:05 am	Presentation on Procuring/Creating High Quality Assessments – Brent Maddin
10:05 am – 10:20 am	Setting the Stage for Ongoing Work
10:20 am – 10:30 am	Transition to State Team Time Rooms

Self-Assessment, Introductions

Where States Stand with Implementation

Assets

Challenges

Review Work Group Plan and Prioritize Deliverables

What deliverables in the plan are our highest priorities?

What deliverables are the lowest priorities?

Is the plan missing anything important?

Where do I see myself and my state getting involved?

Deliverables:

State

- X number of co-developed refined tools/documents
- Quality control system frameworks – populated by proposed and developed tools and processes – to ensure high quality SLOs and assessments

RSN

- A one-page primer of do's and don'ts
- A two-page summary of high quality training practices
- A model quality control system framework that ensures high quality SLOs and assessments
- A protocol and video exemplar for convening evaluators to ensure consistency of SLO quality
- A library of high quality SLOs and supporting tools/materials available through the RSN portal

Supporting Quality SLOs:

What States Can Do

Dr. Leigh McGuigan, TNTTP

Student Learning Objectives (SLOs)

- Classroom-level goals for student learning
- Used to assess teacher performance as one part of a comprehensive evaluation system
- Approaches vary among states

SLO Weights Vary

State	Weight	Number
DC	10%	At least one
INDIANA	10-20%	Primary (all students) Secondary (least prepared)
RHODE ISLAND	50% + (matrix)	Two to four
LOUISIANA	50% (formula)	TBD

SLOs Unpacked: Four Steps

1. Selection of what learning will be measured
(often from among a large and sometimes confusing array of state learning standards, and often in the midst of a transition to Common Core standards)
2. Understanding student starting points
(often where no good data exists)
3. Setting a goal for where all students should end the year
(often without good guidance on where this goal should be)
4. Assessing whether the goal is attained
(often where an assessment must be created)

SLOs are challenging

- **Teachers and principals lack skill, capacity, and tools. Often they have never set a measurable student learning goal – and even this most basic step requires significant skill building.**
- SLOs require a lot of time from teachers and principals, especially in the first year, and especially for teachers with multiple preps.
- Good systems to record and track SLOs do not exist.
- Districts can be reluctant to use existing standardized assessments (cost, perceived lack of alignment to standards, lack of statistical precision), but teacher created and/or administered assessments are often low quality, not comparable across classrooms, and raise concerns about gaming and possible cheating.

Systems can be designed mindfully for improvement over time

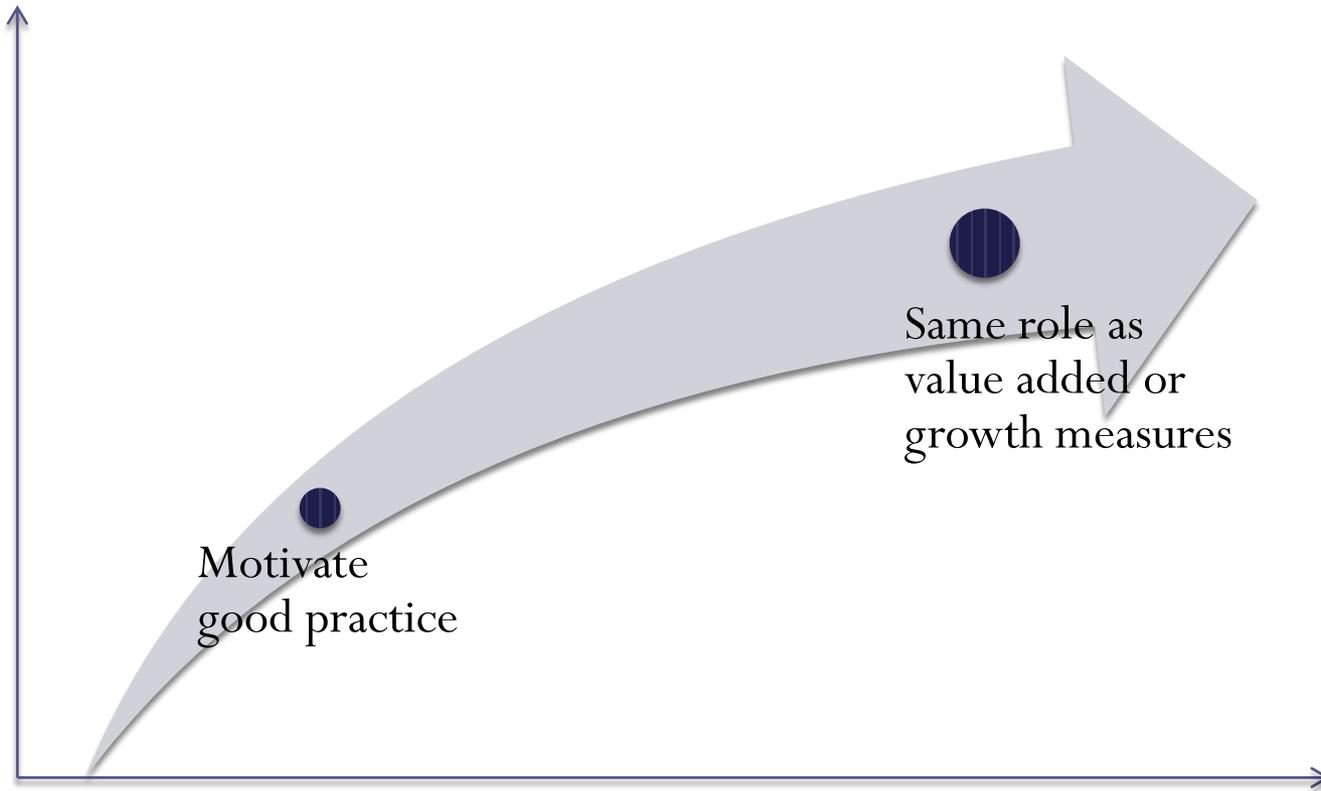
- Where possible, keep weighting moderate (10 to 35%) in initial years and increase as skills, tools and systems improve.
 - This can be accomplished through direct system weights, or by adding in school-wide or other measures in a “50% student learning” system
- Begin with more discretion in scoring
 - E.g. “ OR
..... exceptional growth, typically representing more than 1.5 year of learning
....little growth, typically representing less than a years’ worth of learning ”
- Focus on building skills among district academic managers (Assistant Superintendents, Chief Area Officers) and school administrators

Different theories of action may dictate different approaches to quality control

SLOs motivate good instructional practice	SLOs are substitutes for a statistical measure
<p data-bbox="127 496 915 886"><i>Setting classroom goals for student learning is good instructional practice, promoted by including SLOs as part of the evaluation system</i></p>  <p data-bbox="127 1219 794 1282"><i>Less precision may be required</i></p>	<p data-bbox="983 496 1765 805"><i>In grades and subjects not tested by state tests, SLOs serve as a measure of student learning that create comparability across classrooms</i></p>  <p data-bbox="983 1139 1676 1202"><i>More precision may be required</i></p>

As stakes increase, so does the need for quality control

Weight : Influence on final rating
Consequences: Results for teachers



Accuracy: Accuracy and precision in measurement
Fairness: Comparability across classrooms

Continuum of Quality Control Measures

Provide guidance, templates, tools

Train district staff and administrators

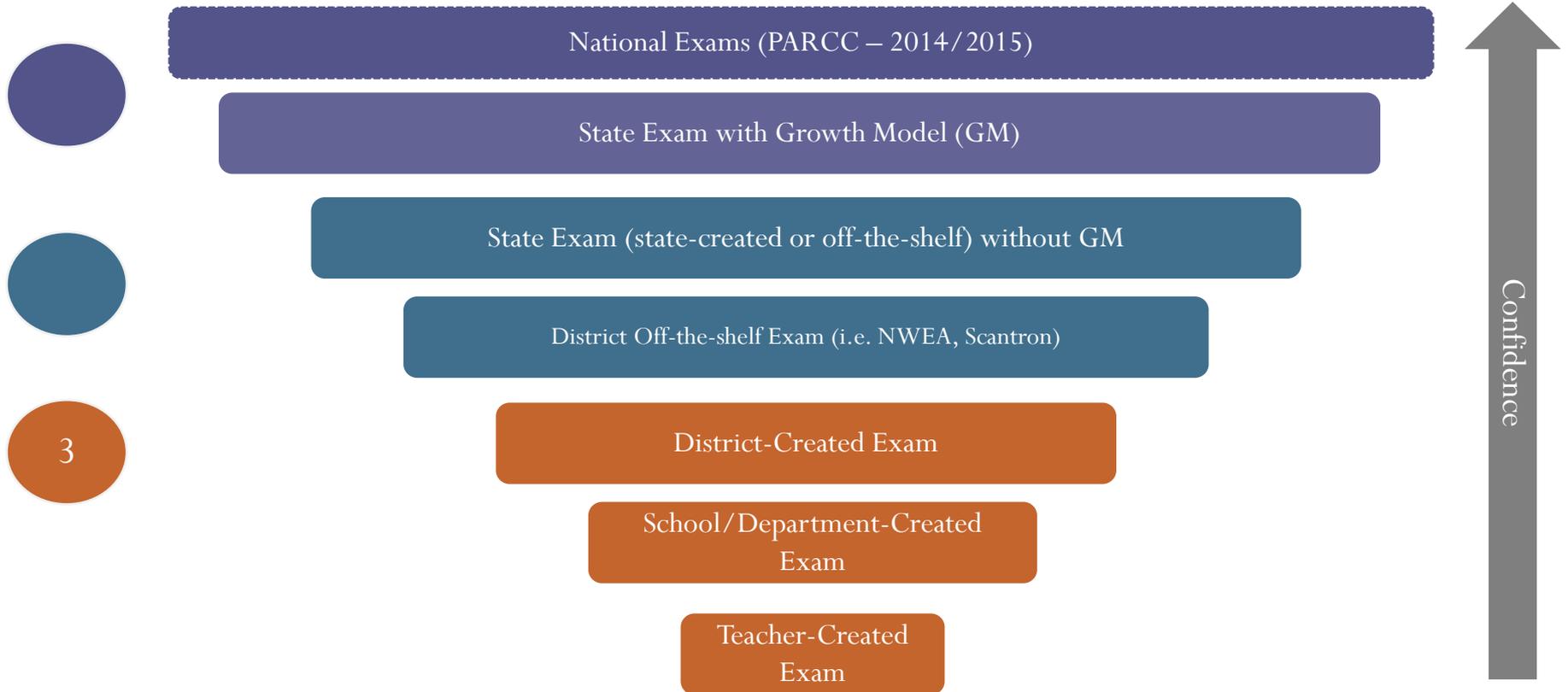
Hold administrators accountable for quality

Audit and impose consequences

Support creation of common assessments (blueprints, consortia, item banks)

Create standardized assessments

The “Assessment Triangle”



3

Pick Any Two

FAST

PRECISE

CHEAP



A Few Resources

- TNTP: Teacher Talent Toolbox
<http://tntp.org/teacher-talent-toolbox/view/evaluation>
- Indiana: www.riseindiana.org
- Washington DC: www.dc.gov/DCPS/impact
- Tennessee: <http://team-tn.org/>
- Rhode Island:
<http://www.ride.ri.gov/educatorquality/educatorevaluation/>
- Delaware:
<http://www.doe.k12.de.us/csa/dpasii/default.shtml>
- New York: <http://usny.nysed.gov/rttt/teachers-leaders/#appr>

Procuring and Creating High Quality Assessments

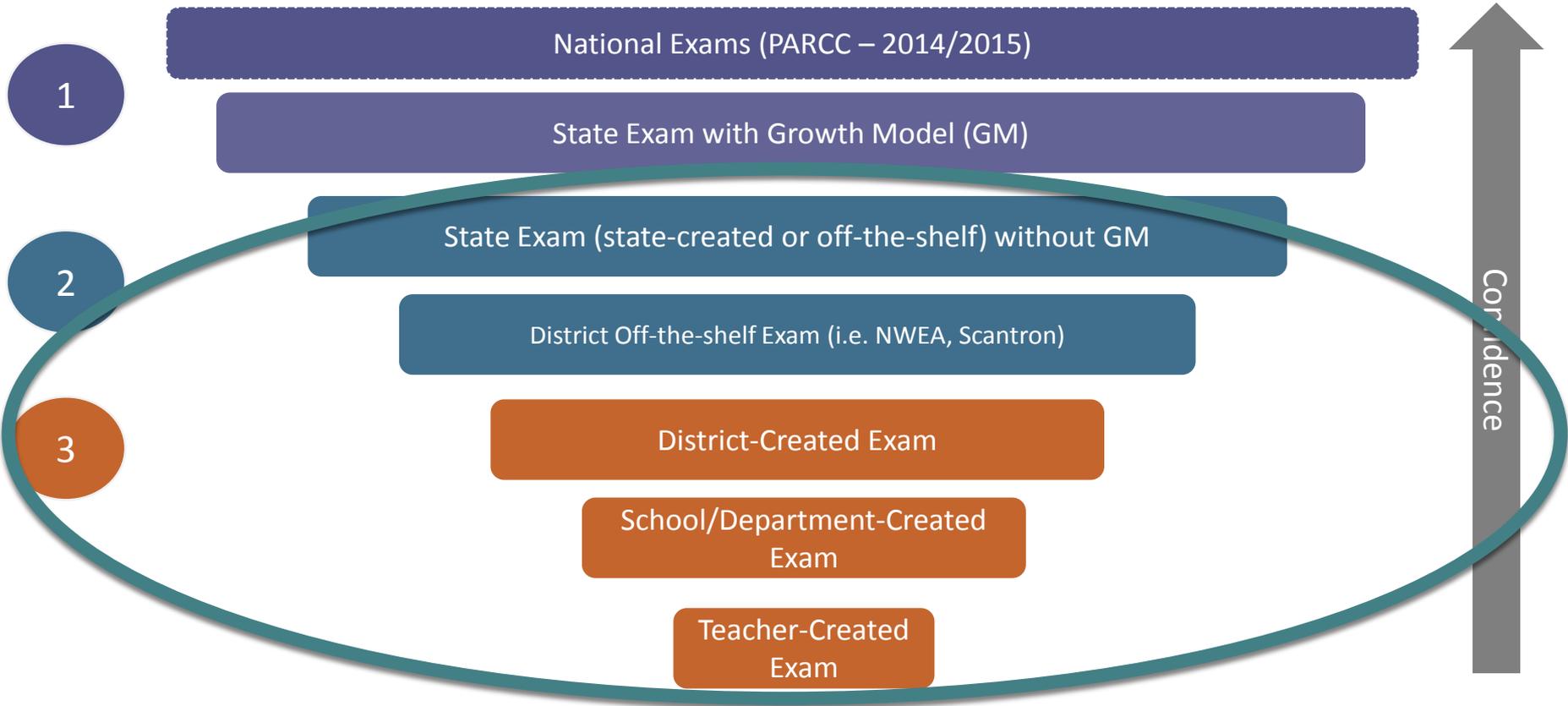
Brent Maddin, Relay Graduate School
of Education

Pro/creating Assessments to Support SLOs

[Procuring + Creating = Pro/creating]

Dr. Brent Maddin

The "Assessment Triangle"



Non-Tested Assessment Challenges

- Non-Tested Electives (PE / Art / Music / etc...)
- Writing
- Test-Friendly Electives (Languages, Econ, etc...)
- 1st – 8th Science / Social Studies
- 9th – 12th Science / Social Studies
- 1st & 2nd Math / ELA
- K – 4 Reading



Challenge

Three Organizing Questions

1. What makes for a high-quality pro/created assessment?
2. How do we assess in *very* non-tested subjects?
3. How can we create high-quality assessments for non-tested grades/subjects at scale?



**1. What makes for a
high-quality
pro/created
assessment?**

[Procuring > Creating]

Procuring Non-Tested Assessment Solutions

Grade/Subject	Possible Assessment Solution
K-4 Reading	<ul style="list-style-type: none">• DRA2 (Grades K-8) <i>(Note, not DRA)</i>• Fountas and Pinnell Benchmark Assessment System (Grades K-8)• Reading A to Z (Grades K-8)• Rigby (Grades K-3)• STEP Reading Assessment (Grades K-3)• Teachers College Reading & Writing Project Reading Assessment (Grades K-8)
1 st & 2 nd Math or ELA	<ul style="list-style-type: none">• ITBS (check alignment)• Stanford 10 (check alignment)• Terra Nova (check alignment)• Item Bank (created, not procured)



Procuring Non-Tested Assessment Solutions

Grade/Subject	Possible Assessment Solution
9 th -12 th Science & SS	<ul style="list-style-type: none">• End-of-Course Exams from Other States (check alignment)• SAT II (relatively high bar, check alignment)• AP / IB (high bar, check alignment)• Item Bank (created, not procured)
1 st -8 th Science & SS	<ul style="list-style-type: none">• Lower Grades:<ul style="list-style-type: none">○ ITBS (check alignment)○ Stanford 10 (check alignment)○ Terra Nova (check alignment)○ Item Bank (created, not procured)• Upper Grades:<ul style="list-style-type: none">○ ITBS/SAT10/TerraNova/ACT (limited alignment)○ Item Bank (created, not procured)

Procuring Non-Tested Assessment Solutions

Grade/Subject	Possible Assessment Solution
<p>Test-Friendly Electives (Languages, Econ, etc...)</p>	<ul style="list-style-type: none">• End-of-Course Exams from Other States (check alignment)• SAT II (relatively high bar, check alignment)• AP / IB (high bar, check alignment)• Item Bank (lower priority)
<p>Writing & Non-Tested Electives (PE, Art, Music, Drama, etc...)</p>	<ul style="list-style-type: none">• No immediate off-the-shelf assessment (UK?)• Bank of Assessments + Rubrics w/ Anchor Work (lower priority)



RGSE: Steps for Creating Assessments

- 1. Teacher identifies standards to be assessed** These are vetted/approved by instructor
- 2. Teacher proposes assessment regime**
 - Interim Assessment Regime (4-6 assessments)
 - Summative Assessment Regime (1-2 assessments)
- 3. Teacher proposes assessment items**
 - Hopefully more procuring than creating
 - Items are vetted/approved by instructor
 - Purposive sample vetted by independent 3rd party

Approving Non-Tested Assessment Items

ALIGNMENT

- Does each item assess the standard it is intended to assess?
- Is there appropriate coverage of standards?
 - Interim Assessments = 100% of proposed standards @ 3+ items/standard
 - Summative Assessments = 50% of proposed standards @ 2+ items/standard

Approving Non-Tested Assessment Items

ITEM QUALITY

- Are all items grade-level appropriate?
- Do items have appropriate weights?
- For multiple choice items:
 - Is there one correct answer?
 - Are the wrong answers true distracters?
 - Do answer choices have parallel construction?
- For constructed response items:
 - Is there a clear/fair scoring guide or rubric?
 - Are there sample student responses?

Writing Non-Tested Assessment Items

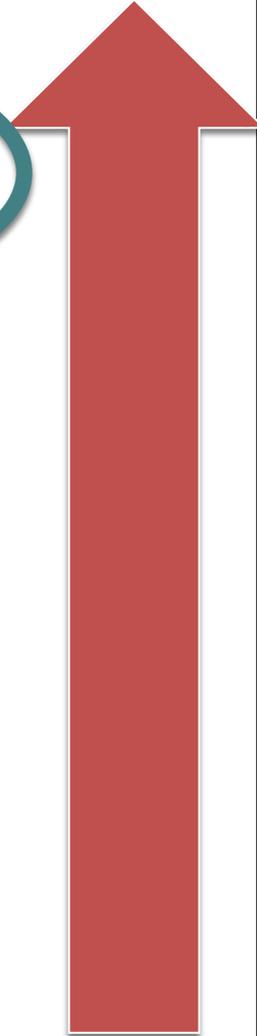
HEADLINE

Having teachers write high-quality assessments is a painful and hard-to-scale solution

2. How do we
assess in *very*
non-tested
subjects?

Non-Tested Assessment Challenges

- Non-Tested Electives (PE / Art / Music / etc...)
- Writing
- Test-Friendly Electives (Languages, Econ, etc...)
- 1st – 8th Science / Social Studies
- 9th – 12th Science / Social Studies
- 1st & 2nd Math / ELA
- K – 4 Reading



Relay Graduate School of Education Model

- **Identify Skills to Assess** (state/national standards)
- **Pro/create Rubric w/ Anchor Work**
 - Create five-point rubric for mastery of skills (or use existing where available—e.g. Six Traits Writing Rubric)
 - Associate anchor work with various levels, when possible
- **Administer Pre/Post Assessments**
 - Determine individual student starting place
 - Set reasonable “on average” growth goal (instructor vets)
 - Administer end-of-year assessment to determine growth



Achievement First Model

Teacher
Career
Pathway



Measuring the Specials Teacher's Impact on Student Achievement

Goal of the AF Teacher Career Pathway: In order to fulfill our mission of closing the achievement gap for our scholars, we have created a process that will help teachers continuously improve and reward exceptional performance: the Achievement First Teacher Career Pathway. There is a high bar of excellence for all teachers in the network and now there is an explicit process to measure excellence and to provide appropriate feedback for teachers to help them progress along the pathway towards excellence.

The Challenge of Assessing Student Achievement for Specials Teachers: All teachers in the AF network will be evaluated using multiple measures, including a teacher's impact on student achievement. Measuring student achievement growth in specials courses can be difficult to do in the absence of state exams, and when achievement is often evidenced in ways other than what can be captured using standardized tests. That is why for "scholar achievement outcomes," the weighting for specials teachers is 20% as opposed to the 40% in subject areas using standardized tests. However, it's important that the process for measuring achievement in specials subjects is rigorous and properly assesses the progress that scholars make in each discipline.

Measuring Excellence in Specials Courses – Scholar Achievement: Three different measures will be used to determine the student outcome portion of the TCP evaluation. The three are listed below with the weighting for each:

- **Skills Assessment – 40%** – Designed to measure individual scholar achievement growth in teacher defined skills measures
- **Group Performance Video – 40%** – Designed to measure group achievement within teacher defined performance criteria
- **Written Knowledge Assessment – 20%** – Designed to measure individual scholar achievement growth in knowledge of the discipline

The teacher has discretion to design each measure as appropriate. The plan for each, including the grading rubrics, methods of evidence collection, skills or knowledge being measured, etc. will be documented on planning templates, and will be presented by the teacher to their school leader and achievement director during the Fall Program Review (page 5).

Soft copies of the templates with instructions on how to complete them, as well as samples of completed assessment templates, are available on the [shared server](#).

In the Spring, evidence from all three measures will be submitted into a portfolio to be evaluated by a review committee during the Spring Portfolio Review (page 6). The three assessment measures are described below.

Achievement First Model

- **Skills Assessment (40%)**
 - ID skills assessed (e.g. 2D drawing, monologues, mile run)
 - Growth from Diagnostic to End-Of-Course on Rubric
- **Group Performance (40%)**
 - ID 2 performances (e.g. exhibition, concert, scrimmage)
 - Articulate performance criteria on a rubric
- **Written Knowledge Assessment (20%)**
 - Students complete Pre/Post Assessment on knowledge of subject (e.g. history, terms, technique, etc...)



3. How can we create
high-quality
assessments for non-
tested grades/subjects
at scale?



In One Word...

COLLABORATION

- Item Bank
- Performance Assessments with Anchor Work



Characteristics of An Effective Item Bank

- Items are independent
 - States/districts/teachers need flexibility
- Items are of various types
 - (MC, CR, Performance Tasks)
- Items are at multiple levels of difficulty
- Items are aligned with CC / state standards
 - w/ "Rosetta Stone" translator between standards
- Slick, web-based interface



Where are These Items?

- Existing State Assessments
- District Assessments
- High-Performing Charter Networks
- Not-For-Profit Organizations (e.g. ANet)
- NAEP, PISA, TIMSS
- ITBS, SAT10, TerraNova, ACT, SAT, SATII, AP, IB
- For-Profit Providers (e.g. testing companies, test prep, text book companies)

Who Should Wrangle the Project?

- Consortium of States?
- Feds?
- Not-For-Profit?
 - New Schools Venture Funded Project/Org
 - Achievement Network

Setting the Stage for Ongoing Work

1. Complete and submit questionnaires to Phil Gonring by noon tomorrow
 - ❖ Identify tools/documents you would like to co-create/refine
 - ❖ Consider volunteering for subgroups
2. Look for revised work group plan and schedule of activities

This information is provided as an example for the viewer's convenience at this convening and is not generated nor endorsed by the U.S. Department of Education. No endorsement by the U.S. Department of Education of any product, commodity, service or enterprise mentioned in these presentations is intended nor should be inferred.

In addition, the instructional practices discussed or shown in the presentations are not intended to mandate, direct, or control a State's local educational agency's or school's specific instructional content, academic achievement system and assessments, curriculum or program of instruction.

Thank You

Please contact SLO Work Group
Facilitator:

pgonring@education-first.org

For problems with Dropbox, contact

Regina Riley:

rriley@education-first.org