New York State
EDUCATION DEPARTMENT

NYSED Evaluation Workgroup

Session #2: Student Performance Measures & Technical Requirements for Inclusion of Assessments in Evaluation

Welcome!

Session #2 Objectives

1. Provide input on the ideal student learning component for educator evaluation

2. Review existing requirements for the student learning component of educator evaluation, and discuss ways to improve that component

3. Process feedback and assess consensus around recommendations

4. Discuss next steps

Agenda

Time	Topic
9:00 am	Welcome and introductions
9:15 am	Reviewing inputs from webinar
9:45 am	The ideal student learning component
11:30 am	Lunch
12:30 pm	How can the consistency in the implementation of SLOs across LEAs and schools be improved?
1:30 pm	How 9:30 pm

Group Norms

- **Presume positive intentions**
- **Fully engage, active listening, and speaking**
- **No cross talk**
- **Respect for everyone's opinions and views. Open to all experiences and views.**
- **Talking piece (something physical to hold)—respect those who speak.**
- **Equal airtime**
- **Respectful of time**
- **No cell phones**
- **Be curious**
- **No need to bash the administration**
- **Ensure all stakeholders' voices are heard**
- **Subgroup work—**

What is our purpose?

- **Provide and capture recommendations to improve educator evaluation system**
  - **Improvements to existing system**
  - **Components and measures of an ideal system**

- **In order to get there we will**
  - **Identify and surface barriers**
  - **Consider technical information provided by experts**
  - **Recommend <u>solutions</u>**

What research tells us about the importance of student outcomes in evaluation

- **Teachers are the single most important school based, and principals are the second most influential factor and have a multiplicative effect related to student outcomes** (McCaffrey, Lockwood, Koretz, & Hamilton, 2003; Rivkin, Hanushek, & Kain, 2000; Rowan, Correnti & Miller, 2002; Wright, Horn, & Sanders, 1997).

- **Students of teachers with higher teacher effectiveness estimates outperformed students of teachers with lower teacher effectiveness estimates** (Cantrell and Kane, 2013).

- **Students assigned to more effective teachers are more likely to attend college, attend higher- ranked colleges, earn higher salaries, live in higher SES neighborhoods, and save more for retirement** (Chetty, Friedman and Rockoff, 2011)

New York State EDUCATION DEPARTMENT

Reviewing Inputs from Webinar #1

9:15-9:45 am

Barriers to an ideal system…

- **Student Learning**
  - Assessments may not fully capture students' progress
  - Year to year changes in assessments make it difficult to understand student growth
  - Factors outside of the classroom teachers' control must be taken into account
  - SLOs are not always implemented to improve teaching practice or student performance
  - Teachers may not have enough knowledge about students at the start of the school year to set useful SLO targets
  - Teachers in schools with high student turnover may develop targets at the beginning of the year that aren't relevant to student in their classes at the end of the year

The Ideal Student Learning Component

9:45-11:30 am

Table Team Activity

Premise: You and your table team will represent a hypothetical LEA, which has been given local control over the student learning component of an evaluation system for teachers and principals.

Guiding Question: What would your ideal student learning component look like?

New York State

EDUCATION DEPARTMENT

How can the consistency in the implementation of SLOs across LEAs and schools be improved?
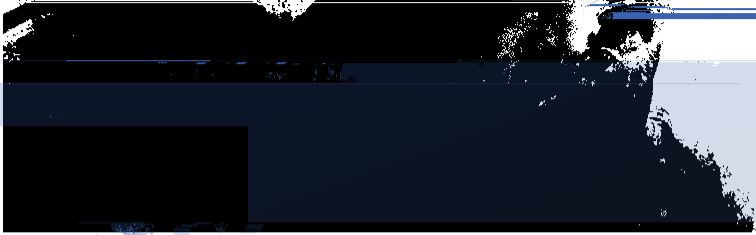**The SLO Process**

12:30-1:30am

A Little History

1983

- 25 states include a definition of SLO's in their teacher evaluation systems

- SLO's for individual teachers in 23 states
- SLO's for teams of teachers or grade levels in 3 states, optional in 7
- Schoolwide SLO's required in 3 states and optional in 4 states

Common Elements of State Definitions

Element	Number of States
Measurable	12
Based on student growth and achievement	16
Aligned with state or local standards	9
Based on prior student learning data	9
Measure teacher impact on student learning	4
Aligned with course content	4

Assessments Used to Evaluate Student Learning Objectives

Assessment type or feature	Number of states
National or state standardized assessment	14
District-wide or school-wide measures	12
Classroom-based measures	12
Test Vendor-developed content	3
Comparable across classrooms	5
Valid and reliable	3
Aligned with state standards	2
Rigorous	2

Why use SLO's?

- SLO process contains key aspects of good instruction: review of student data, goal setting, progress monitoring, reflection

-

Basic SLO Process

Review Student Data → Develop SLO → SLO Approval → Monitor Progress → Score SLO

Implementation Strategies to Help Ensure Consistency and Rigor

- Provide exemplary SLOs across subject areas

- Approve assessments for use in SLOs

- Assessment literacy training

- Build Principal capacity to assess and provide feedback to improve SLO quality and rigor

- School or team-based goals (individual targets)

- Mid-year SLO review

- Student data use training

- Randomly sample SLOs for audit

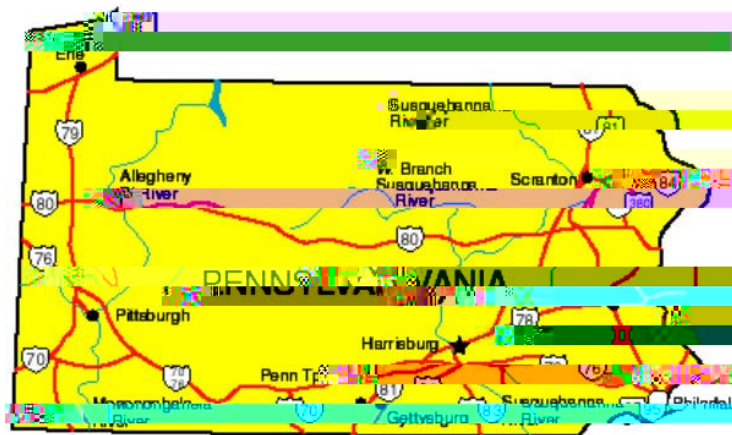- Consideration of SLO quality/rigor in scoring SLO
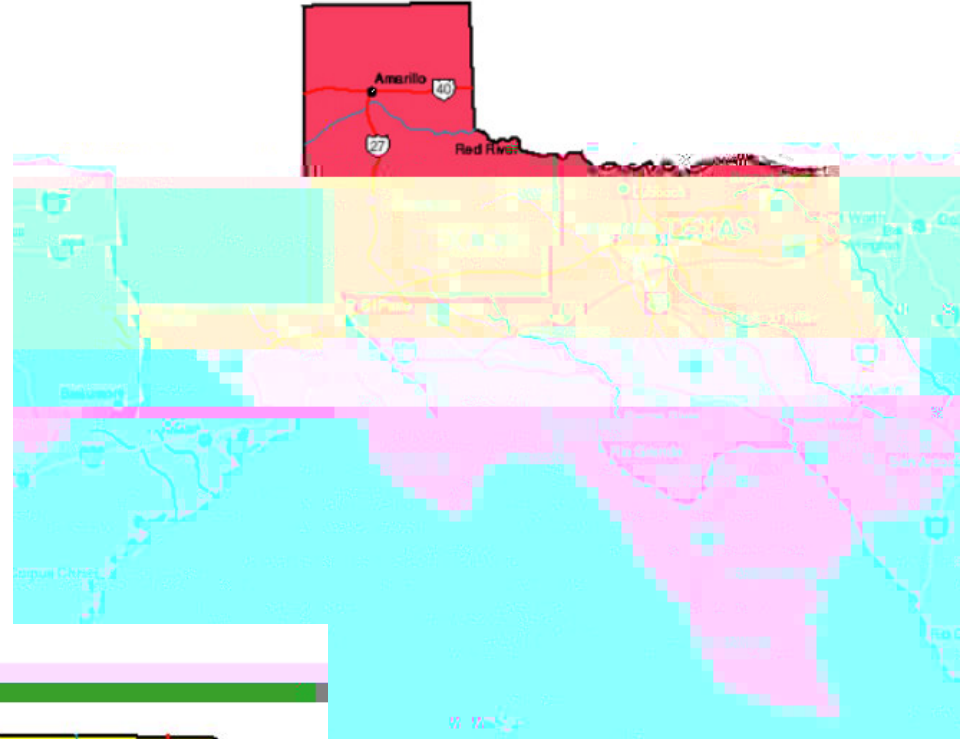
Examples From Other States

Table Talk #1 (25 minutes)

Group Discussion #1 (20 min)

- **Please share out your or your group's decisions and discuss the rationale.**

How can the assessment quality be balanced with inclusion of additional assessments in SLOs?

1:30-2:15 p.m.

Assessments in SLOs

- **Goal: have assessment choices available for all educators to use in SLOs with (a) high degree of ownership/buy-in; and (b) sufficient technical quality**

- **Current Reality:**

    - **Some grades/subjects have more (and better) assessments than others**

    - **Many teachers dislike the use of traditional standardized assessments in SLOs and prefer locally-developed options**

    - **Locally-developed and classroom assessments have greater buy-in but more technical challenges**

    - **Educator capacity around assessment development is often low, although can be built over the long term**

Desired Technical Properties of Assessments (abridged)

Validity

- **Validity: the degree to which evidence and theory support the interpretation of test scores for their proposed uses**

- **Not a single statistic; an ongoing process**

  - **Documented alignment to content standards**

  - **Involvement of educators in item design and review**

  - **Varied set of items by level of cognitive complexity and item type**

Reliability

- **Reliability = precision/stability of results**
- **Would student scores change if:**

    They got a different set of items that purported to measure the same knowledge?

    Someone else scored their assessments?

    They took the same test another time?

Fairness

- **High-quality assessments must enable ALL students to demonstrate their knowledge (UDL principles):**

  Precisely-defined constructs

  Clear instructions

  Maximum readability

  Allowable accommodations for SwD and ELL

  Items free of bias (DIF analysis)

Educator Capacity: Assessment Development

-

Table Talk #2 (15 min)

Guiding Question: When considering the use of locally-developed and third-party assessments in educator evaluator systems, how can the State balance the need for assessments that meet certain technical criteria with the desire for LEAs to have flexibility in their approaches to evaluation?

Group Discussion #2 (15 min)

- **Please share out your or your group's recommendations and discuss the rationale.**

Agenda

Student Growth Measures

2:15 -2:45 pm

Statistical Growth Models (In Brief)

- A **group of models** **d**esigned to measure the

Selected Observations from NY Data

- Concern: student growth is influenced by factors that educators don't control (creating potential disincentives)

- Data show low correlations between classroom factors (poverty, SpEd, etc.) and SGPs; this means there is very little "penalty" for teaching these kinds of students (as would be the case with proficiency rates)
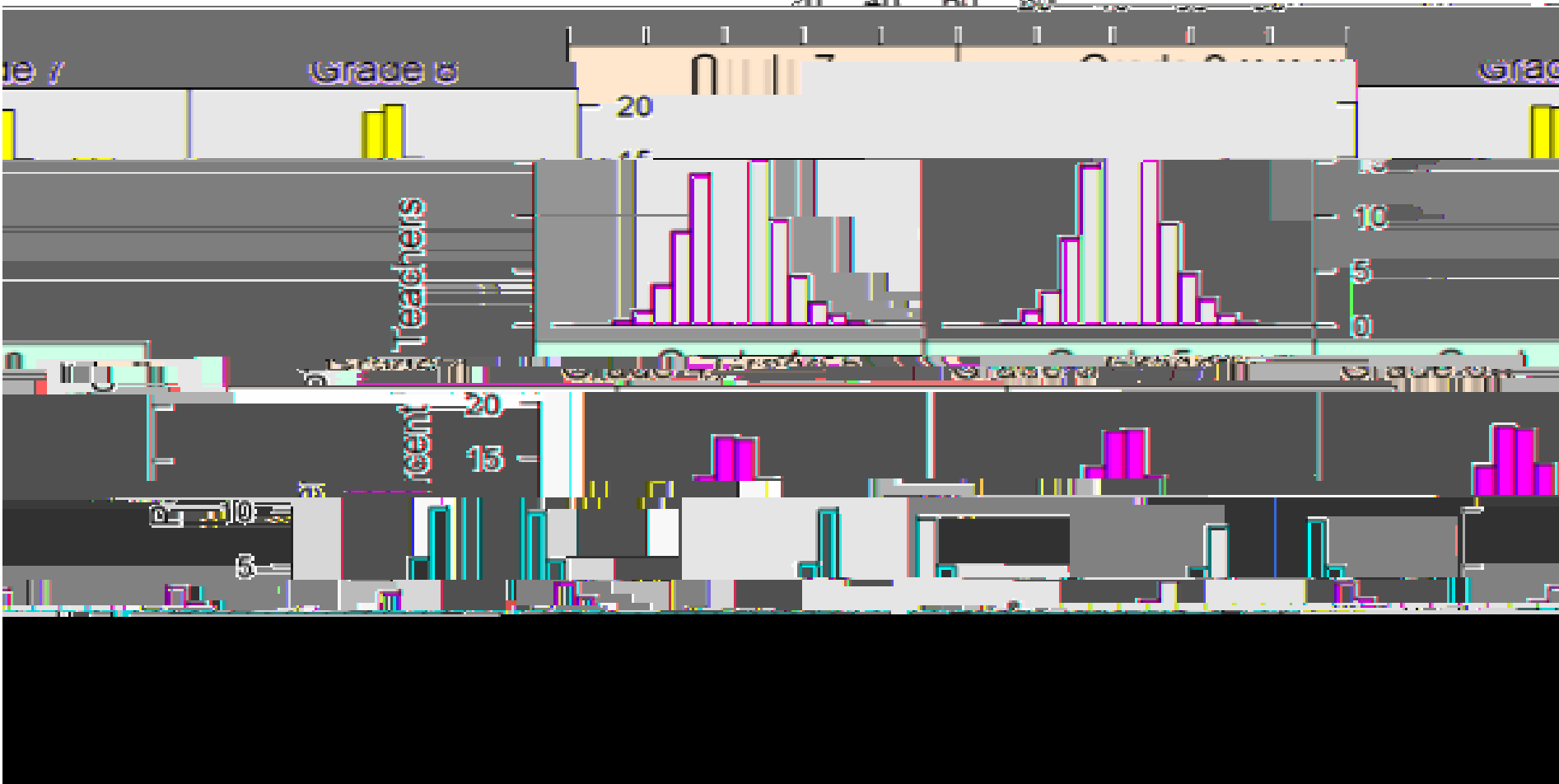
Selected Observations (cont'd)

- Concern: distribution of Student Growth scores is fundamentally different from other measures
- Data confirm this to be true, although most educators are still Effective or Highly Effective on Student Growth

Figure 5. Distribution of Grades 4–8 Teacher MGPs by Grade, Adjusted Model
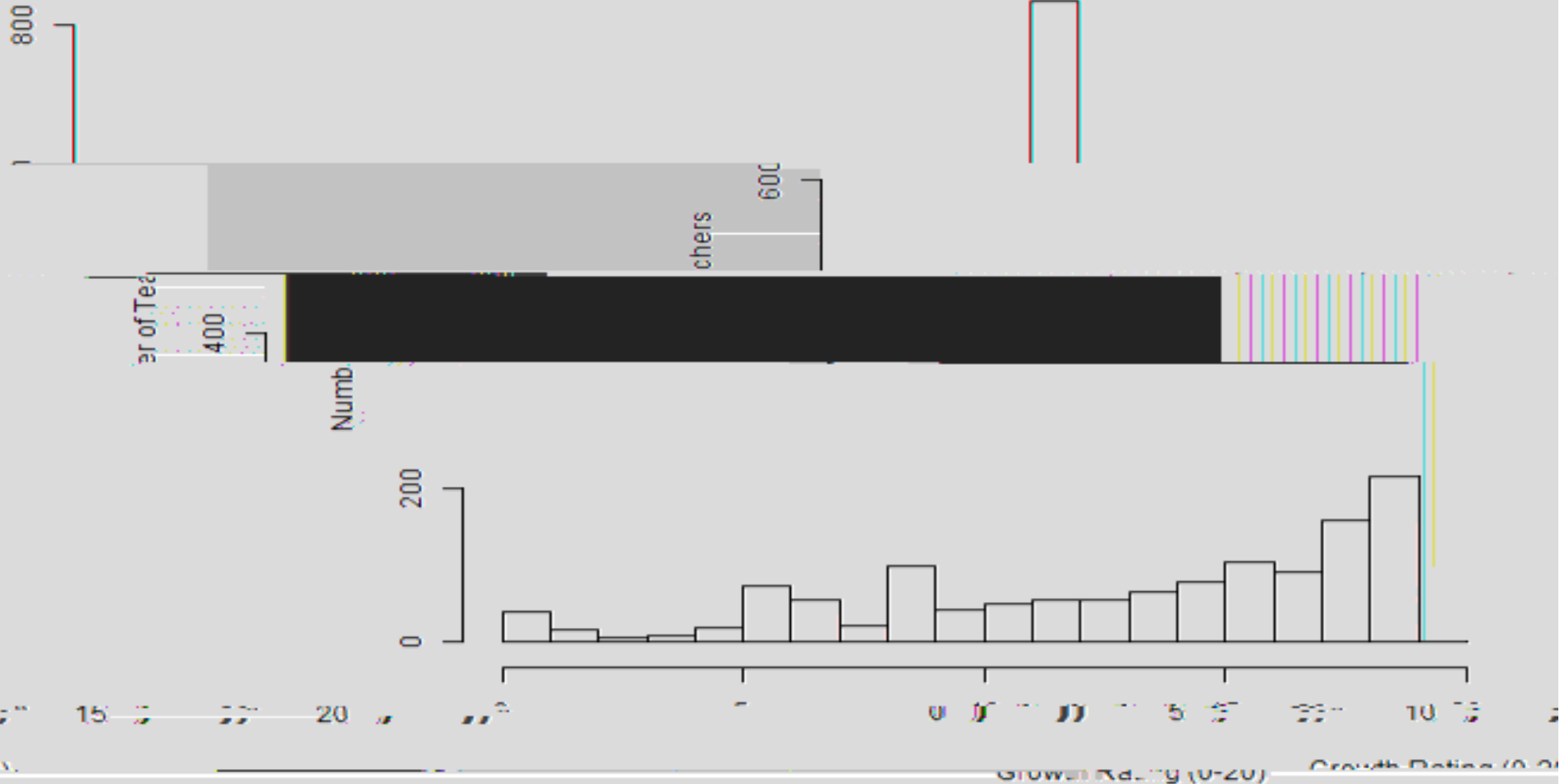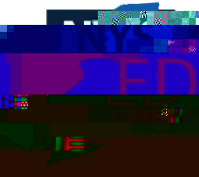
Growth on Comparable Measures

Grades 4–8 (including, for instance, schools ... 16.

schools, and principals who serve students in G... serving Grades 4–12) from 2012–13 to 2015–1...

...chool, and Principal Growth Ratings

Table 15. Grades 4–8 Teacher, S...

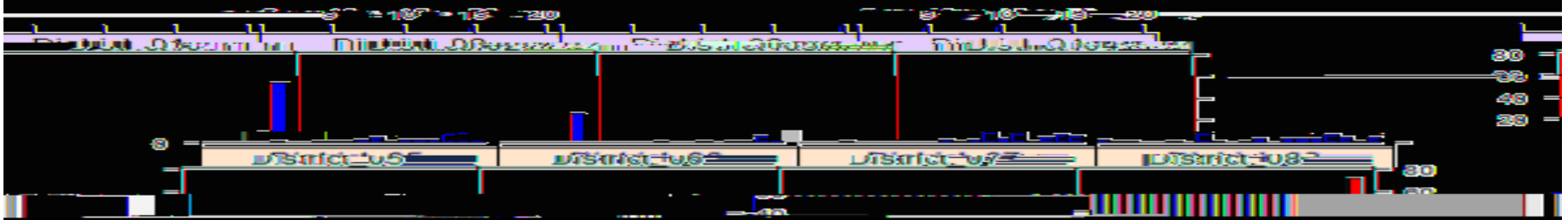| School Year	Level	Highly Effective	Effective	Developing	Ineffective
	School	9%	75%	9%	7%
	Teacher	8%	77%	10%	6%
13–14	Principal	6%	77%	10%	7%
	School	7%	76%	10%	7%
	Teacher	7%	77%	11%	5%
14–15	Principal	7%	77%	10%	6%
	School	7%	76%	11%	6%
	Teacher	8%	76%	11%	5%
15–16	Principal	7%	77%	10%	6%

Growth on Comparable Measures by District

Table Talk #3 (15 min)

Consider the following two scenarios:

•

Guiding Questions

- How does each scenario address identified barriers?

- What concerns do you have with each scenario?

- Is there one scenario that your group recommends?

Agenda

Time	Topic
9:00 am	Welcome and introductions
9:15 am	Reviewing inputs from Webinar #1
9:45 am	The ideal student learning component
11:30 am	Lunch
12:30 pm	How can the consistency in the implementation of SLOs across LEAs and schools be improved?
1:30 pm	How can the assessment quality be balanced with inclusion of additional assessments in SLOs?
2:15 pm	Student Growth Measures
2:45 pm	**Closing & Next Steps**

Closing

- Next steps: summarizing workgroup recommendations for student learning component of evaluation system

- Next topic: Educator practice component and other measures

- Next webinar: February 14th, 3:30-5:00pm

- Next in-person meeting: March 7th