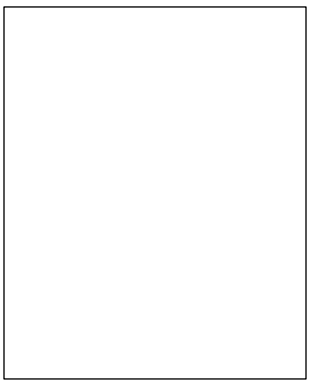
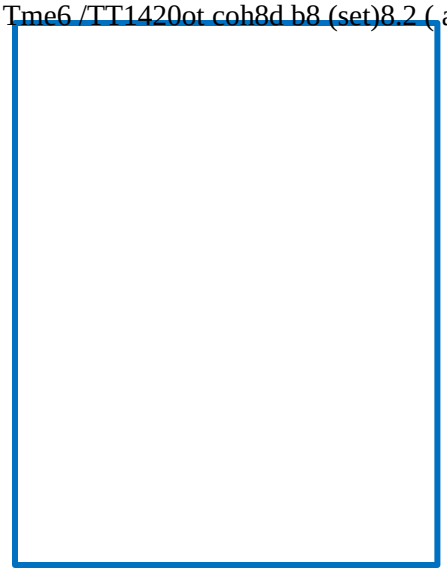
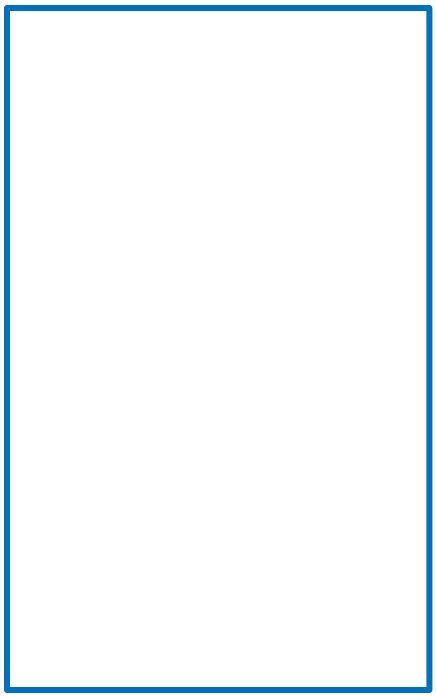
Session 3: The third session took place on January 10, 2018. This all-day, in-person session was designed to focus on student performance measures and technical requirements for the inclusion of assessments in evaluation. Major objectives for session 3 were to receive input from participants on the ideal student learning component for educator evaluation and to gather recommendations on ways to improve the



help them reflect on practice. Workgroup participants also noted that SLOs present challenges for districts with high student turnover, since the cohort for which the goals are set at the start of the year is likely to change significantly by the end of the year.





APPR Timeline

Timeline

New York State's Evaluation System (cont.)

2014-15:

- Governor signs Chapter 56 of the Laws of 2015, establishing a revised evaluation system for teachers and principals (Education Law §3012-d).
- All districts are required to have an approved APPR plan under the new statute by November 15, 2015 or to have an approved Hardship Waiver.

2015-16:

- 18% (n=122) of districts have approved plans under Education Law §3012-d; 82% (n=567) remain under Education Law §3012-c with an approved Hardship Waiver.
- All districts must have an approved APPR plan under Education Law §3012-d by December 31, 2016.
- At its December meeting, the Board of Regents adopts a transition period during which time the results of the grades 3-8 ELA/math State assessments and any State-provided growth scores are to be used for advisory purposes only. Separate transition evaluations that exclude these measures will be provided to affected educators.

2016-17:

-

Education Law §3012-d APPR Components

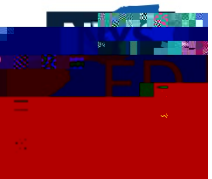
Education Law §3012-d

Components of the APPR Evaluation System

Overview of APPR Transition Period Regulations (2015-16 through 2018- 19 school years)

The Governor's Common Core Task Force "unanimously affirms the importance of adopting and maintaining high educational standards and rigorous performance measures to increase the competitive standing of, and therefore the opportunities for, all our students."

Richard Parsons, Chair, Common Core Task Force



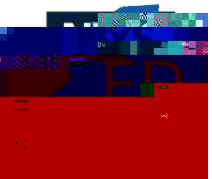
Many of the recommendations in the Task Force report reflect areas the Board and Department have discussed and are taking action on.

Session #1 Objectives

1. Share roles of the NYSED Evaluation Workgroup
2. Provide feedback and recommendations on purpose and use of evaluation system
3. Understand intended purpose and use of evaluation
4. Discuss and give feedback on each component of evaluation system to inform future meetings
5.

Agenda

Time	Activity
9:00	Welcome and introductions
9:45	Consider purposes and uses of evaluation
10:30	Reflect on experiences with evaluation and identify focus areas for future meetings
11:30	Lunch
12:30	Discuss personalized PD and data sources
2:45	Share next steps

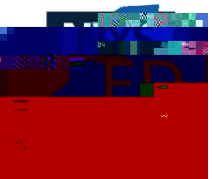


Agenda

Time	Activity
9:00	Welcome and introductions
9:45	Consider purposes and uses of evaluation
10:30	Reflect on experiences with evaluation and identify focus areas for future meetings
11:30	Lunch
12:30	Discuss personalized PD and data sources
2:45	

Partner Conversations and Share Out (12 min)

- In your ideal evaluation system – conceptually, what would be the most important purpose? What would be the most important use?



Current Evaluation System (8 min)

Two major purposes

1. Employment decisions
2. Teacher and principal development

Current Evaluation System

- What does the legislation say?
  - € Evaluation system = Student performance + Observation
    - Student performance – Must use state-provided growth score, if available; otherwise, student learning objective
      - Note optional second subcomponent may be used
    - Observation – Conducted by supervisor and trained evaluator external to the school
      - Note option to use trained peer rated effective or higher
  - € Weighting and scoring ranges of components – Must be transparent at the beginning on the school year
  - € Matrix – Must be used to determine composite score
  - € Prohibited elements –

Group Discussion (10 min)

- In your ideal evaluation system – what would you add or change about:
  - The overall system purpose
  - The use of data from the system
  - The system itself (components, processes, etc.)

  - Note: Future discussions and recommendations will be grounded by the shared agreement related to these

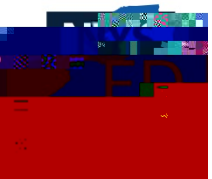
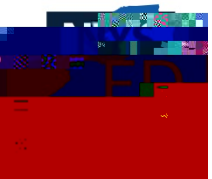


Table Talk Part 2 (15 min)

What issues or decision points would you like the Evaluation Workgroup to continue to discuss and provide recommendations on related to the evaluation system's use of STUDENT PERFORMANCE?

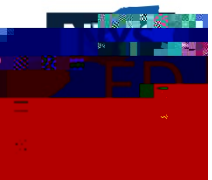
Consider your thoughts on:

- State-provided growth scores
- Student learning objectives
- Locally selected measures of student growth



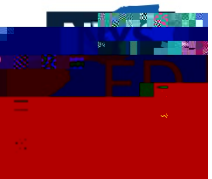
Let’s Share (15 min)

- Share one issue or decision point that your table discussed to have the Evaluation Workgroup continue to collaborate on and provide recommendations on related to the evaluation system’s use of
  - Observations, and
  - Student performance



Agenda

Time	Activity
9:00	Welcome and introductions
9:45	Consider purposes and uses of evaluation
10:30	Reflect on experiences with evaluation and identify focus areas for future meetings
11:30	Lunch
12:30	Discuss personalized PD and data sources
2:45	Share next steps




Personalized Profiles

Learning Outcomes

- Increase participant's understanding of personalized professional development plans.
-

Overview Of Personalized Professional Learning Plan

Research shows that when it is well designed and well implemented, professional learning has the power to strengthen practice and improve student learning.

January 2017 | A New Vision for Professional Learning Learning Forward Educat ionCounsel

Questions To Consider

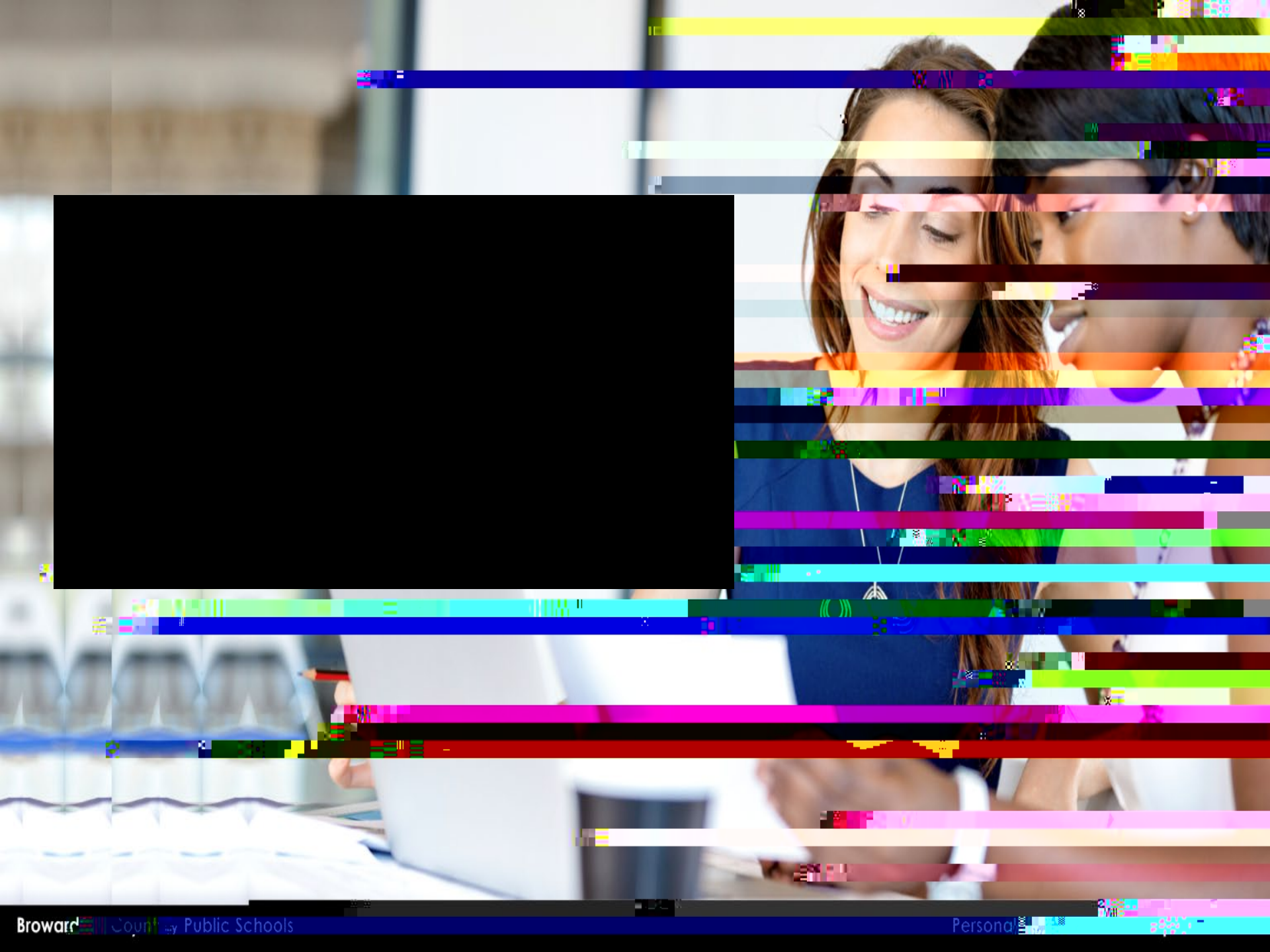
ESSENTIAL QUESTIONS FOR PROFESSIONAL DEVELOPMENT:

1. Why is professional development important?
2.

Questions To Consider

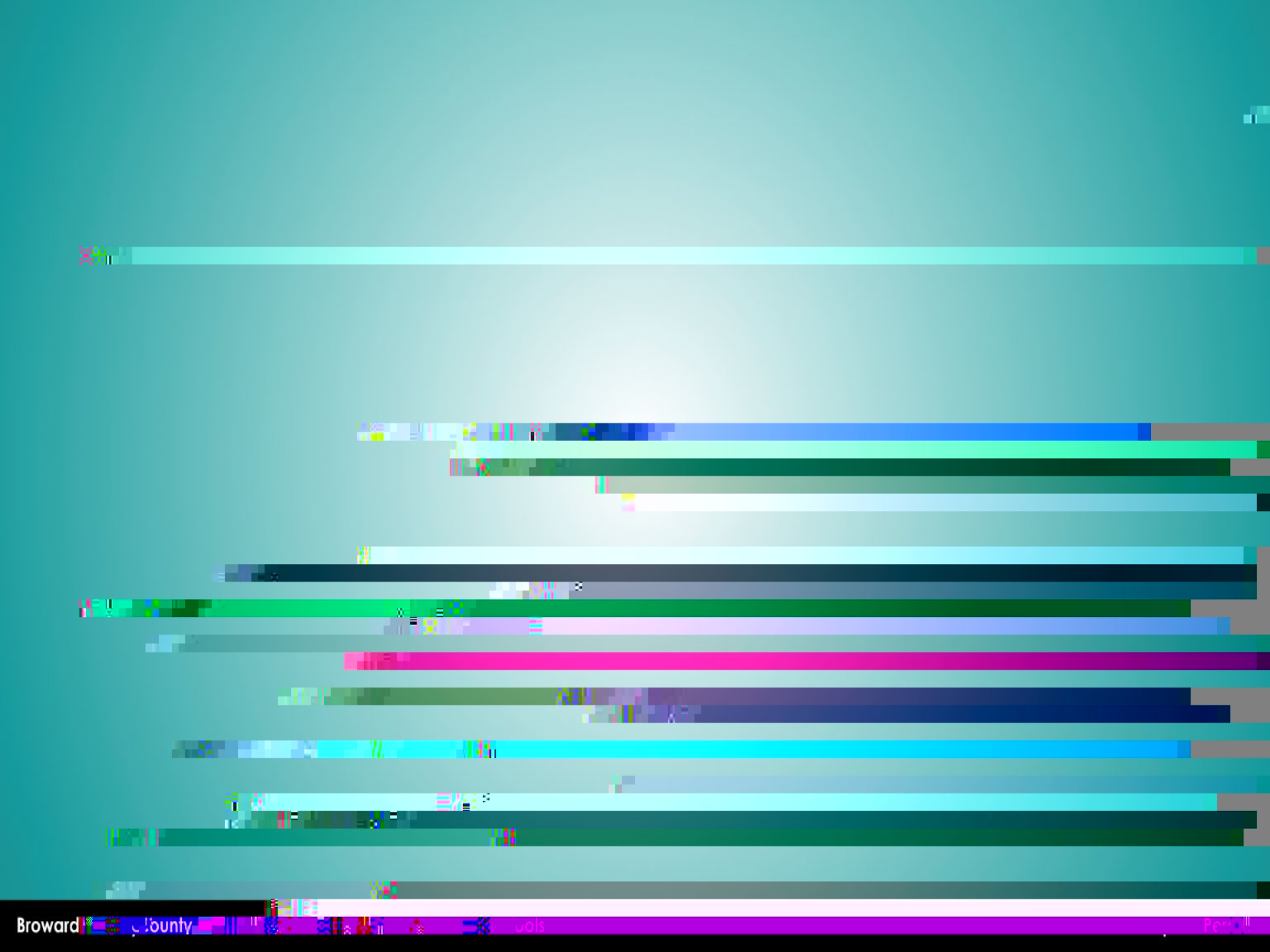
ESSENTIAL QUESTIONS FOR PROFESSIONAL DEVELOPMENT:

5. How do we continue to learn and grow to keep adapting and evolving in the professional development arena for the adult learners?
6. How do we ensure that all professional development is based on the needs of the whole child?
7. How do we ensure that professional development aligns with the goals of the district and schools?



Broward County Public Schools

Personal


Broward County

What It’s Not

- One shot professional development
- One shoe fits all
- Evaluative
- An “I got you”

Embedded Vidt7<I

What The Research Says

Enables educators to directly explore, discover, and adapt strategies needed to meet their individual needs.

It allows educators to practice new instructional strategies, and apply knew knowledge immediately in a classroom - based, or embedded learning environment, to address adult learning needs.

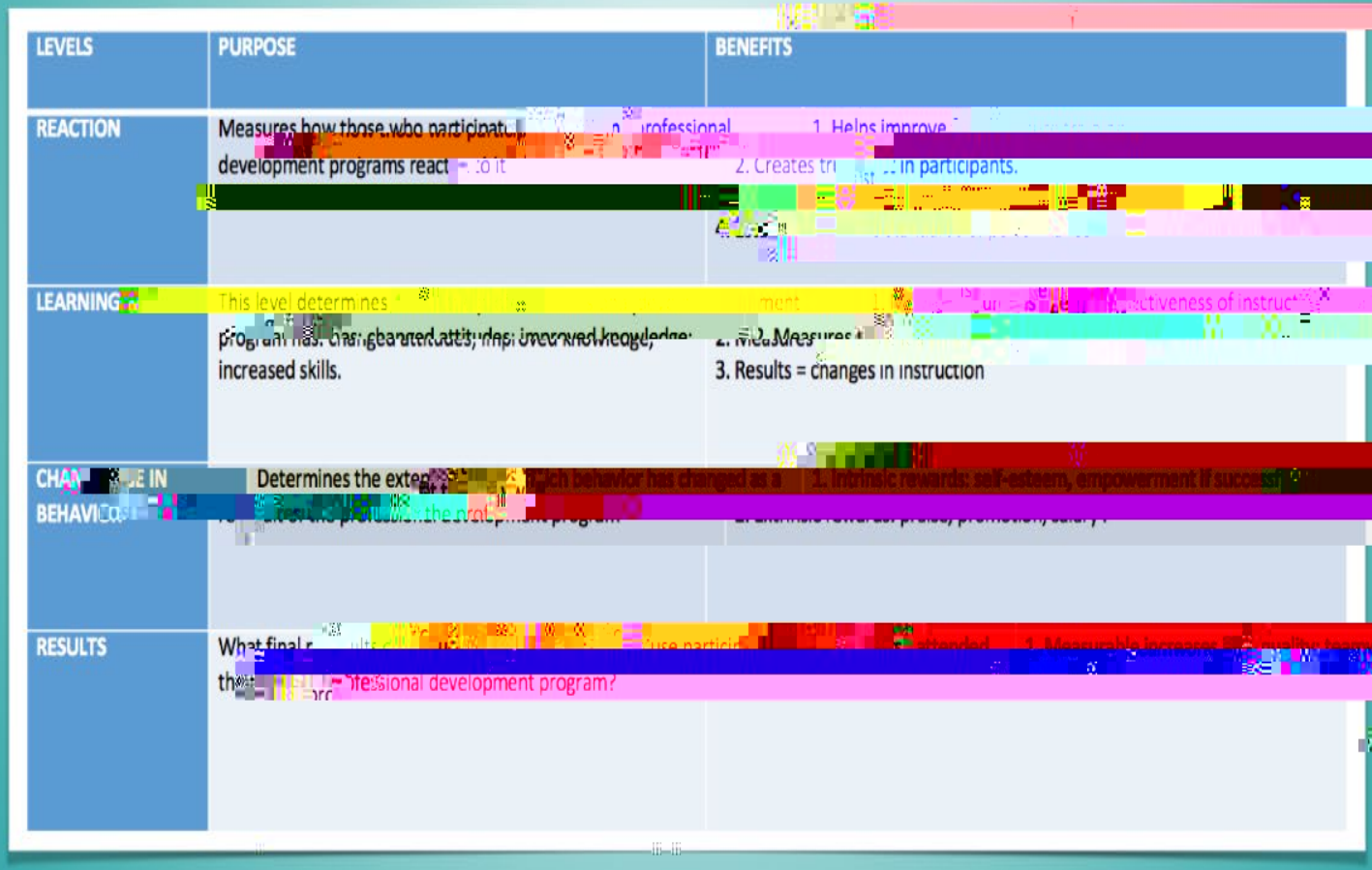
"What is Personalized professional Learning?"
By KristiMeeuwse and Diane Mason June 28, 2017

Advice To Adult Learners

- Set a cooperative learning climate.
- Create mechanisms for mutual planning.
- Arrange for a diagnosis of learner needs and interests8 (for m)4 (ut)4 (ua)-4 (l p)-4 (la213 T<</w)n8Tf 0 Tw -0.904 -1.976 Td <0083>Tj EMC /LEnnl
-

TT Au6 (T1-3l6 3( )-3ID 257)T8En2(n)du6 (-3c)-706e .6 (T8on8 2[( )i4 3( )-3o4.94 7n0 3)1.

Levels for Consideration when Evaluating Personalized Professional Development Plans



LEVELS	PURPOSE	BENEFITS
REACTION	Measures how those who participat[illegible] [illegible]rofessional development programs react [illegible]:o it	1. Helps improve [illegible] 2. Creates tri[illegible]: in participants.
LEARNING	This level determines [illegible]ment [illegible]ctiveness of instruct[illegible] pro[illegible]am has changed attitudes, improved knowledge, increased skills.	2. Measures [illegible] 3. Results = changes in instruction
CHA[illegible]E IN BEHAVIO[illegible]	Determines the exte[illegible]ich behavior has changed as a [illegible]the profitable the profitable program.	1. Intrinsic rewards: self-esteem, empowerment if succes[illegible] 2. Extrinsic rewards: praise, promotion, salary[illegible]
RESULTS	What final r[illegible]use particip[illegible]attended [illegible] th[illegible]rofessional development program?	1. Measurable increases in quality, teamwo[illegible]safety[illegible]

Using Data To Inform Personalized Professional Learning Plans

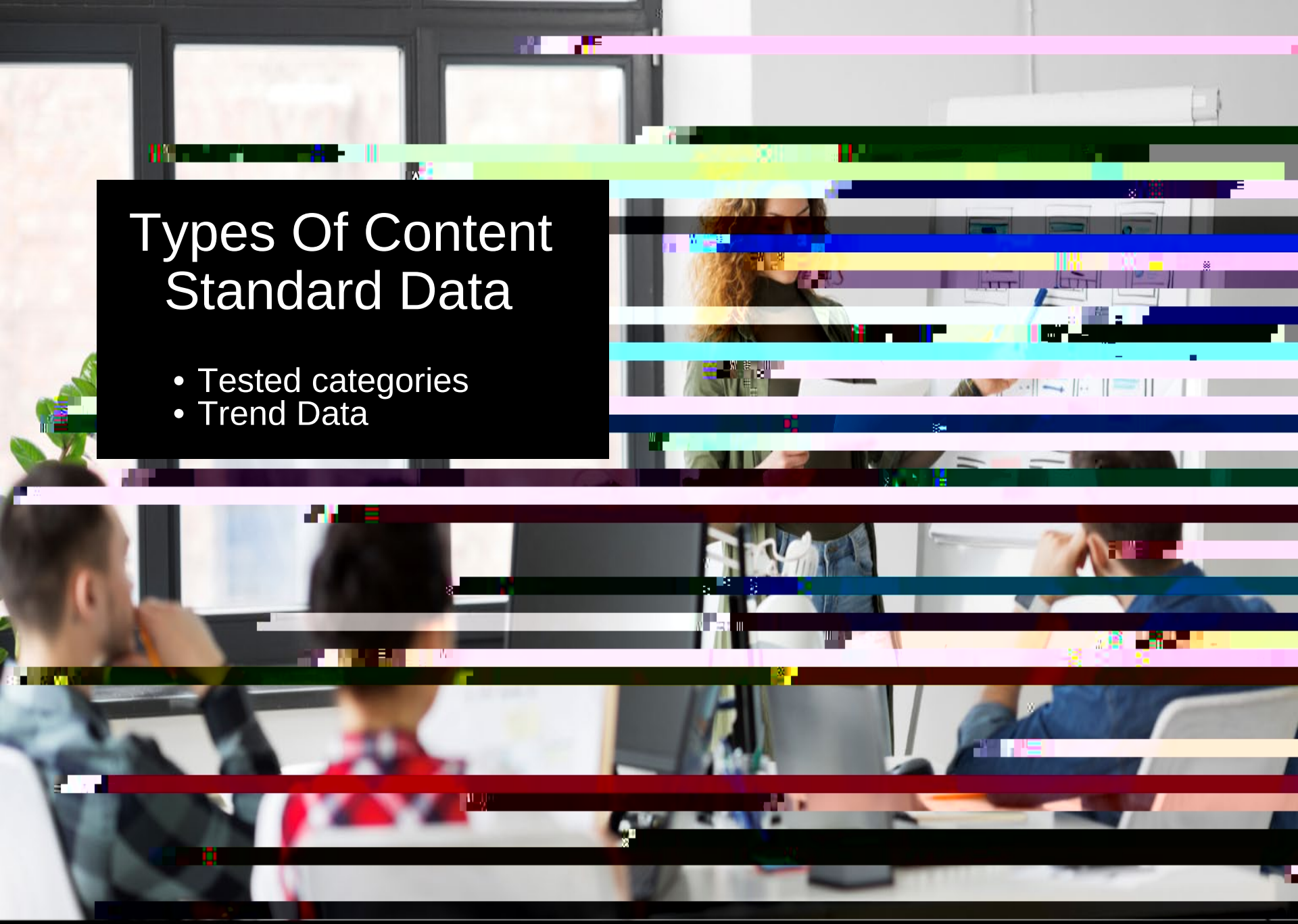
Data- driven includes using a variety of sources and types of student, educator, and system data to identify learning needs, set goals, plan, assess, and evaluate professional learning, preferably in a cycle of ongoing learning and improvement.

Types Of Student Data

- • • St1720 2pes 1DApes 1D

Types Of Instructional Practice Data

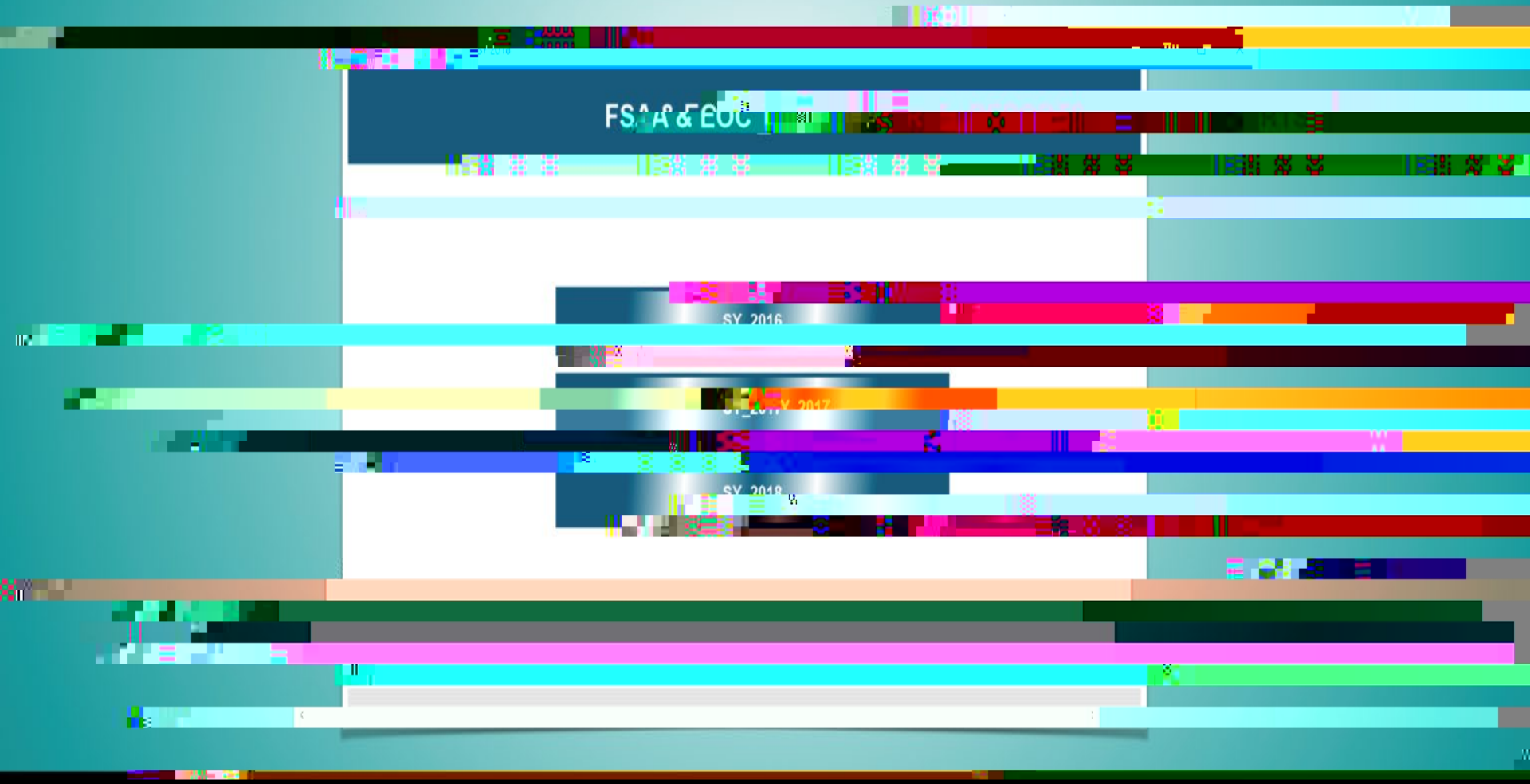
- School-wide growth
- Deliberate practice
- Results based on teacher performance
- Specific topics within the evaluation rubric



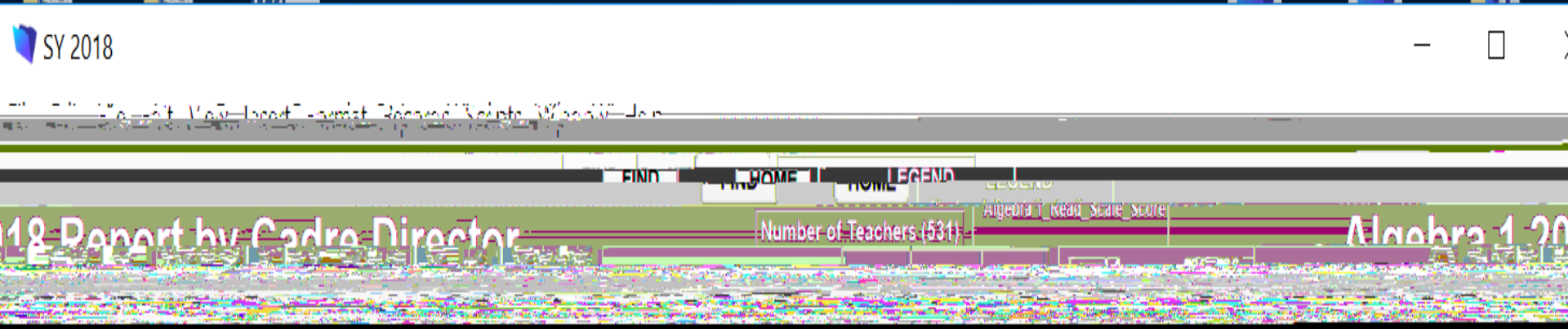
Types Of Content Standard Data

- Tested categories
- Trend Data

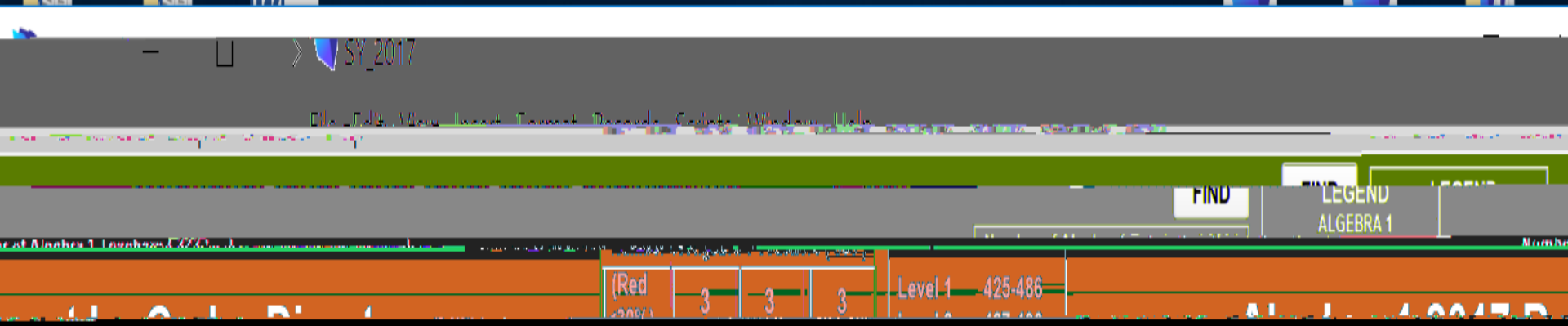
Using Data To Drive Personalized Learning Plans



Professional Development Plans



Professional Development Plans

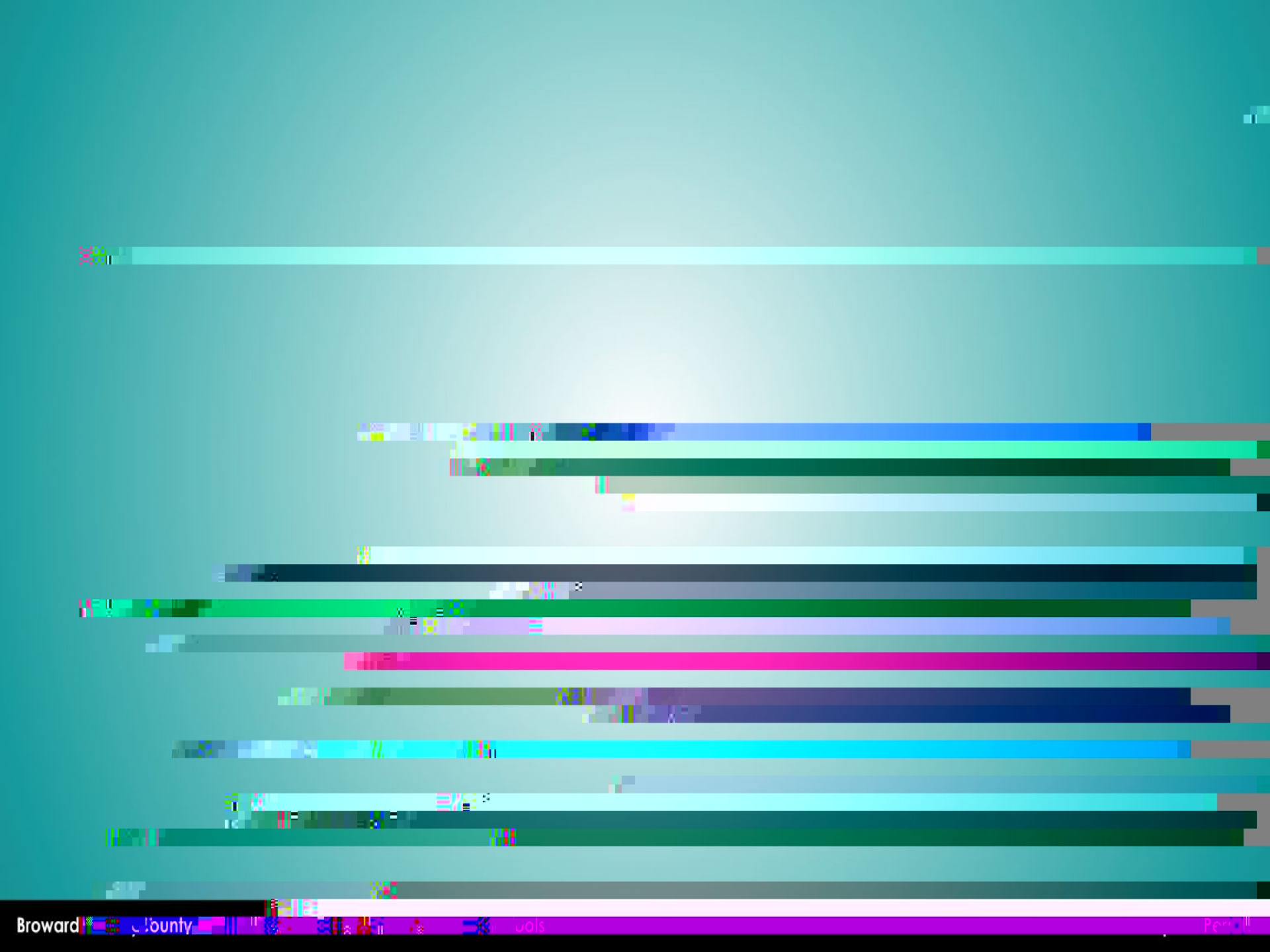


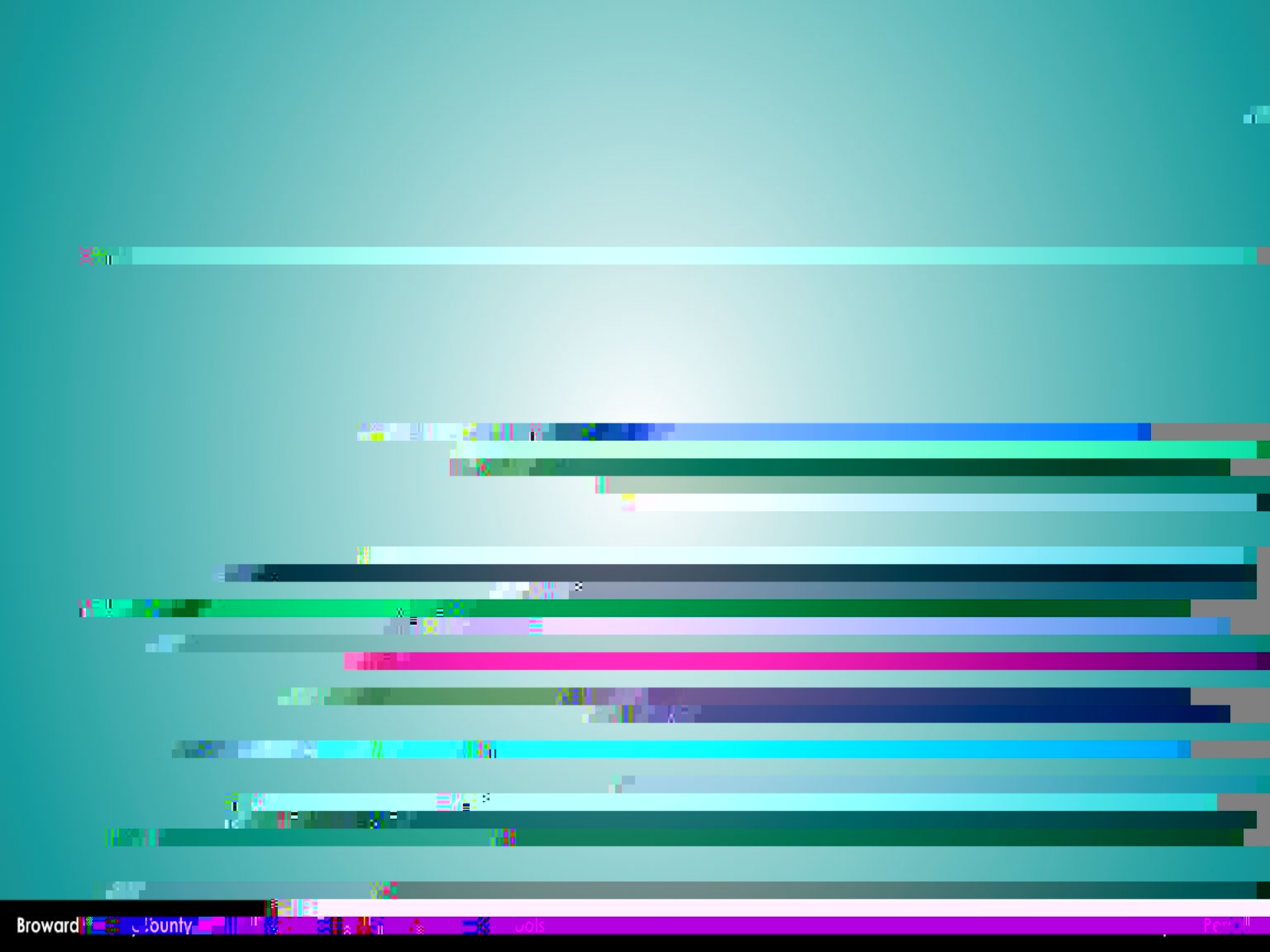
Professional Development Plans

Professional Development Plans

Potential Elements for Personalized Professional Development Plans

- Timeline for achieving growth
- Mannerin which growth will be assessed
- Activitiesto support growth in identified areas
- Artifactsthat serve as benchmarks of growth


Broward County


Broward County

Appendix B

Summary of 2015 Evaluation Changes

Evaluation – Summary of Changes

- The enacted budget created requirements and options for a new evaluation system administered by the Department in accordance with Commissioner's Regulations promulgated by the Board.
- The new evaluation system is comprised of two components that determine each educator's rating:
  - Student performance: Requires the use of a state-provided growth score, if available; otherwise requires the use of a student learning objective (SLO). SLOs must use State assessments, as available.
    - If added by local collective bargaining, an optional second subcomponent could be used, comprised of an additional state-provided growth score on a state test or a growth score from a state-designed supplemental assessment calculated using a state-provided or approved growth model. These state-designed supplemental assessments include those developed, designed, purchased, or acquired by SED.
  - Observations: Requires observations by a supervisor and an independent evaluator from outside the school building.
    - Districts also have the option of having observations conducted by a trained peer who has been rated Effective or Highly Effective.

Appendix C

Uses of Annual Eval0.5 (0.5 7.7317p])5 72(i)6 7

What are annual evaluations used for?

According to the NY law:

Evaluations shall be a SIGNIFICANT FACTOR for EMPLOYMENT DECISIONS including but not limited to:

- promotion,
- retention,
- tenure determination,
- termination, and
- supplemental compensation.

Such evaluations shall also be a SIGNIFICANT FACTOR for TEACHER AND PRINCIPAL DEVELOPMENT, including but not limited to:

- coaching,
- induction support, and
- differentiated professional development.

Appendix D

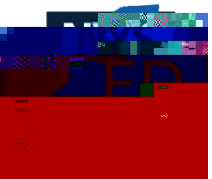
Evaluation Workgroup Session 2

NYSED Evaluation Workgroup

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

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

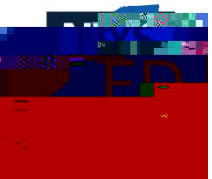


What is our purpose?

- Provide and capture recommendations to improve educator evaluation system

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



Reviewing Inputs from Webinar #1

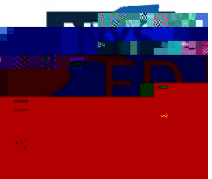
9:15-9:45 am

An ideal evaluation system….

-

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



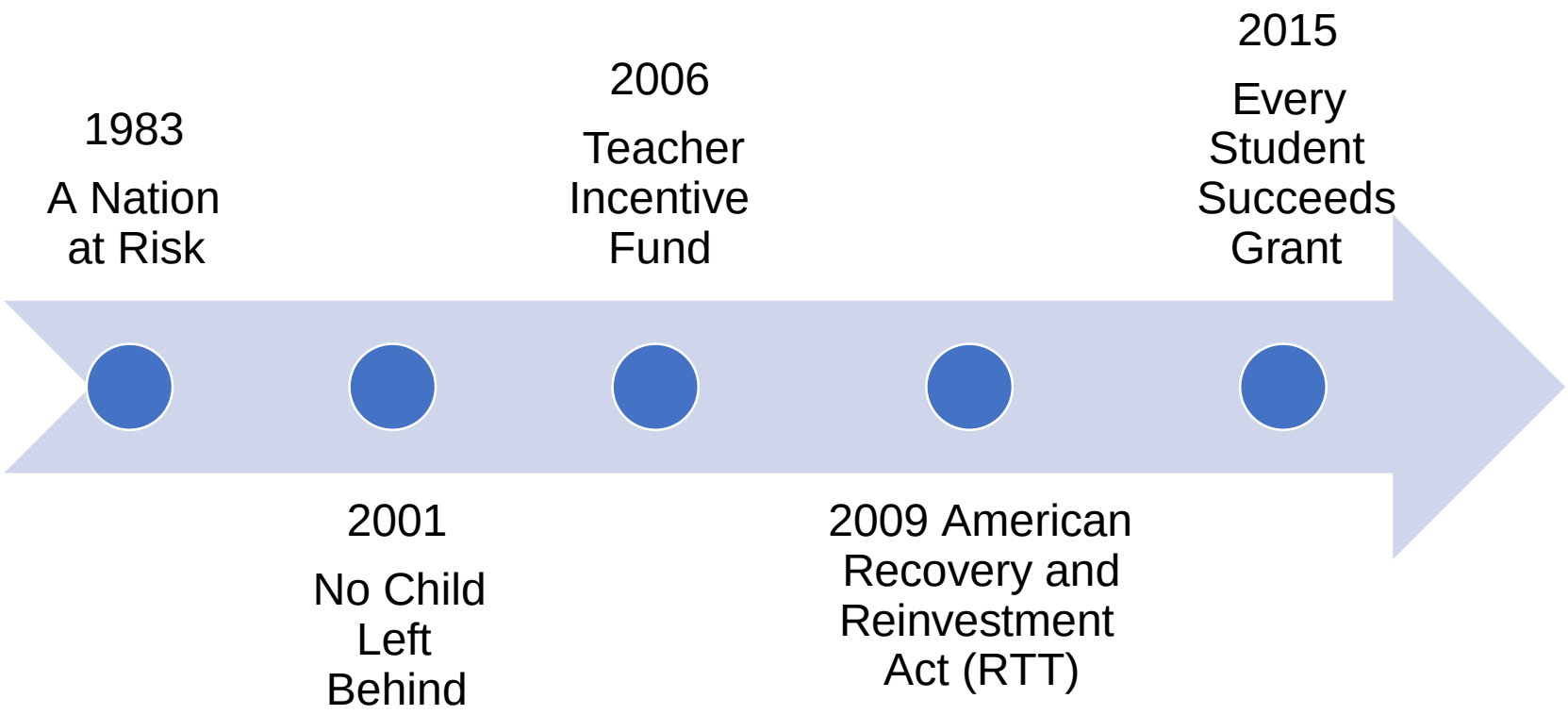
The Ideal Student Learning Component

9:45-11:30 am

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

12:30-1:30am

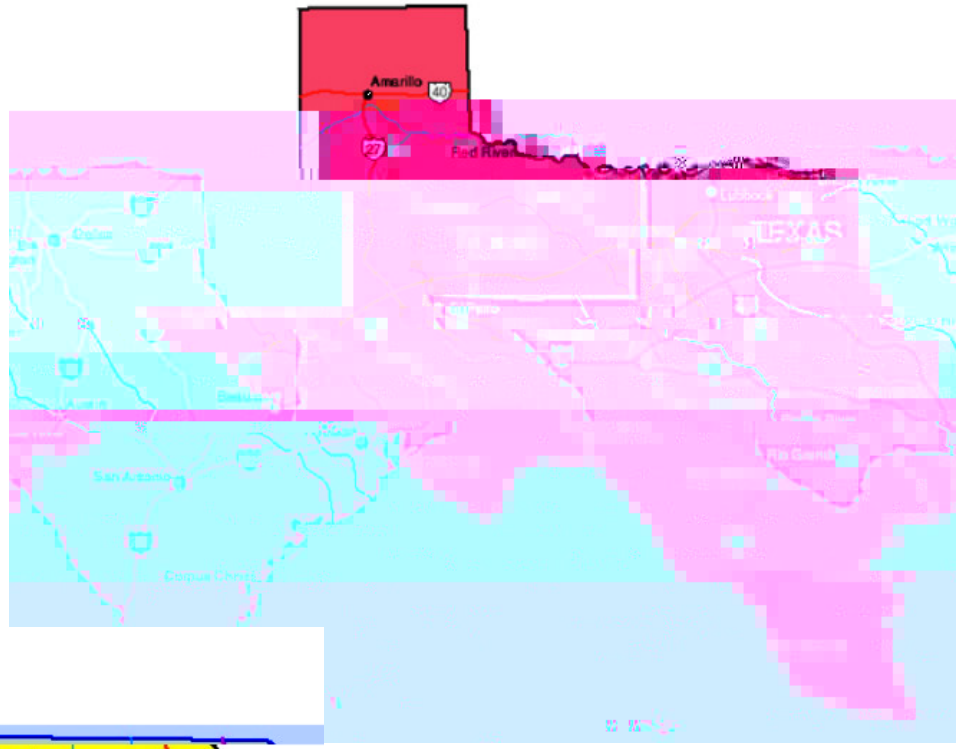
A Little History



- SLO’s for individual teachers in 23 states
- SLO’s for teams of teachers or grade levels in 3 states,

Basic SLO Process

Examples From Other States



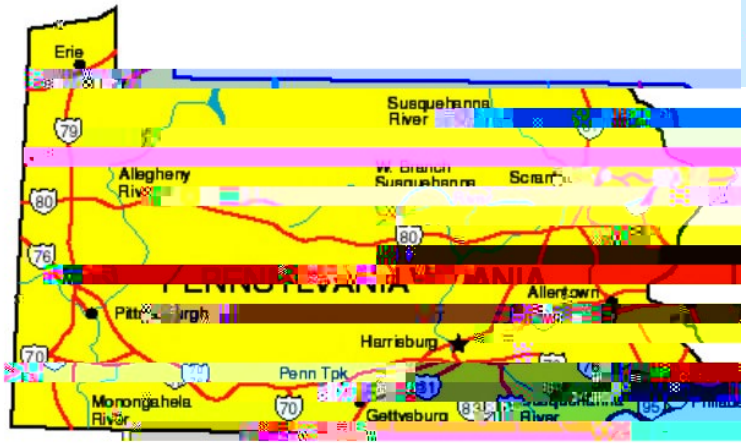
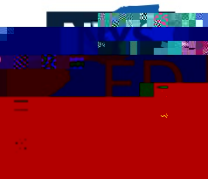


Table Talk #1 (25 minutes)

Guiding Question: Thinking about the SLO l C (S)-4J ET Q (ng
ung1 (e)-2 (tha)9 (t 45A5yC (S)-2 (u w)10 ()-2 (ul)-1 (d r)1)A5e

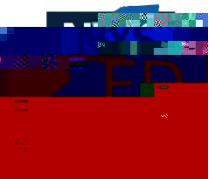


Assessments in SLOs

- 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
  - €

Desired Technical Properties of Assessments (abridged)

- What are some key aspects of assessment quality and why do we have them?
- From the Standards for Educational and Psychological Testing (2014 edition):
- Part I: Foundations
  - Validity
  - Reliability/Precision and Errors of Measurement
  - Fairness in Testing
- Part II: Operations
  - Test Design and Development
  - Scores, Scales, Norms, Score Linking, & Cut Scores
  - Test Administration, Scoring, and Reporting

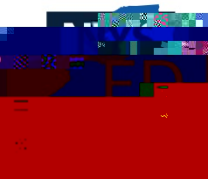


Validity

- Validity: the degree to which evidence and theory

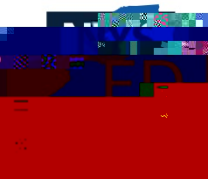
Reliability

- 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)



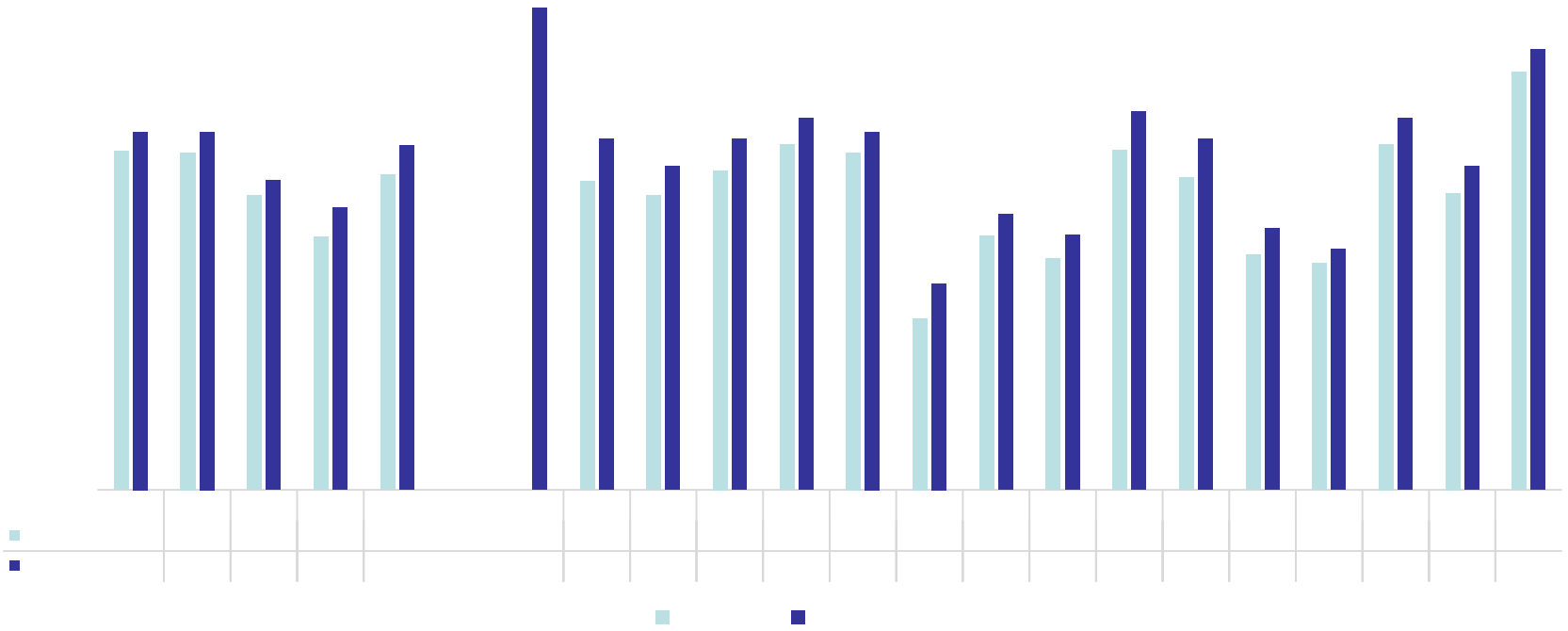
Test Design

- Is the test scaled or simple # correct?
- Under a pre/post arrangement (typical for many SLOs), are pre and post equated for difficulty?
- If cut scores exist, how were they established?
- Does the assessment contain enough items to accurately differentiate student knowledge (are there items for low, medium, and high performers)?
- What are reasonable (and ambitious) expectations for growth? How much do they vary based on students' starting point?

Educator Capacity: Assessment Development

- Most educators get relatively little training in this area …

FfT Ra640.9 (t)-7.8 (re)-0.9 (l)2W 5.88 -122.76 re f 202.32 255.48 61/ 0 1 26.1s8 1



Key Decision for States/Districts

- Many (most?) educators prefer to use locally-developed or classroom assessments for SLOs.
- However, many of these assessments have low (or unknown) technical qualities, and improving educator capacity in this area is a long-term project.
- In the short term, states must weigh the tradeoffs of greater educator buy-in (from a more flexible approach to allowable assessments and how growth targets are set) vs. ensuring minimal technical quality (from approved assessment lists and pre-determined growth target formulas).

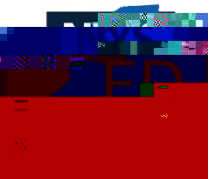
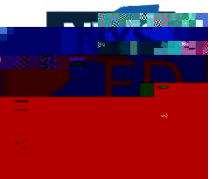


Table Talk #2 (15 min)

Guiding Question: When considering the use of locally

Agenda

Time	Topic
9:00 am	



Student Growth Measures

2:15 -2:45 pm

Statistical Growth Models (In Brief)

- A group of models designed to measure the contribution of schooling at various levels (school, grade, classroom, etc.) to gains in student performance over time.
- Uses statistical techniques to separate the impact of schooling from other factors that may influence growth, but are generally beyond the control of schools/educators (prior achievement, EcDis, SpEd, ELL).
- Goal: provide information on what different levels of

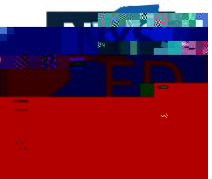
Selected Observations from NY Data

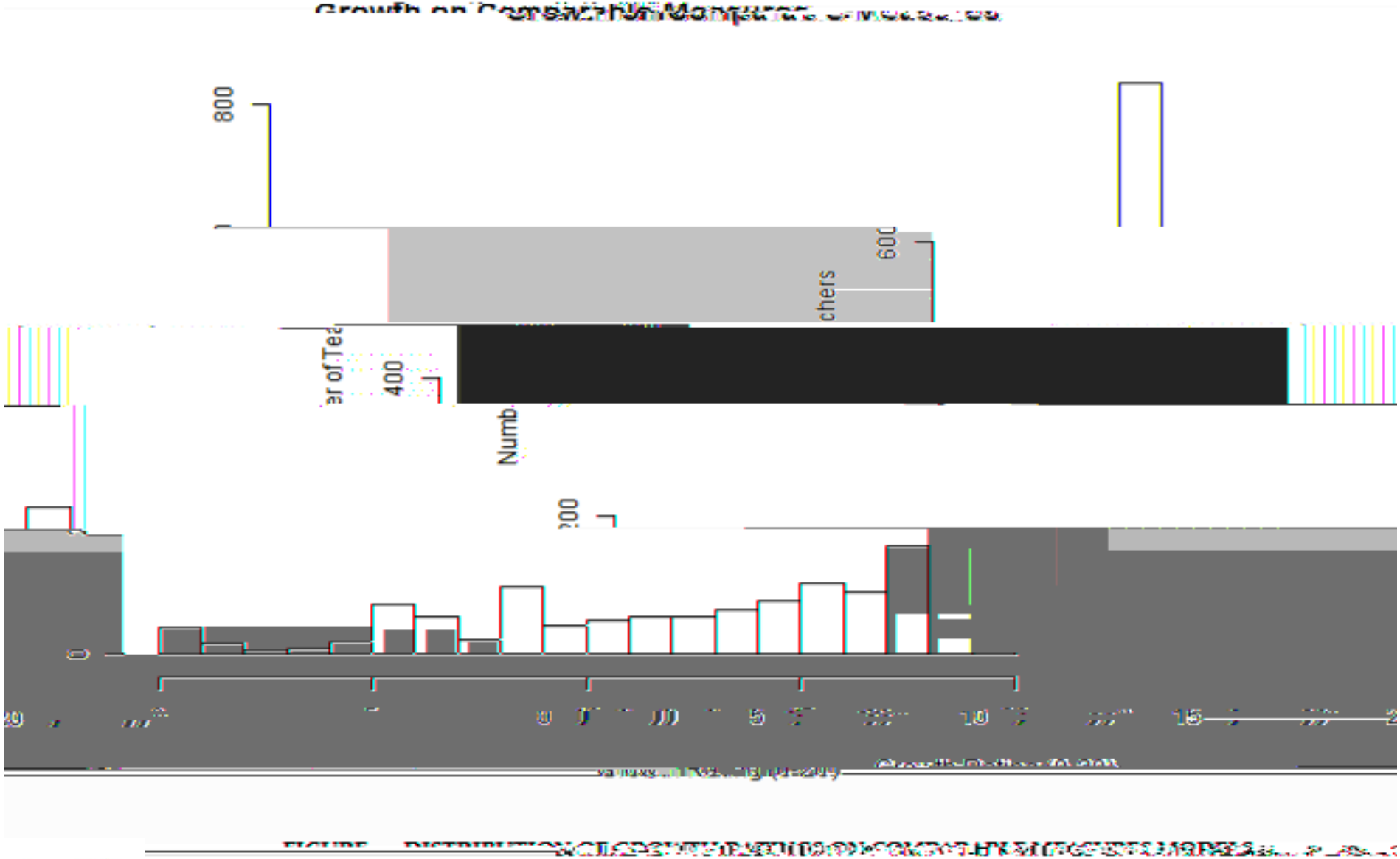
- Concern: student growth is influenced by factors that educators don’t control (creating potential disincentives)
-


100

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




800
600
400
chers
Numb
0
5
10
15

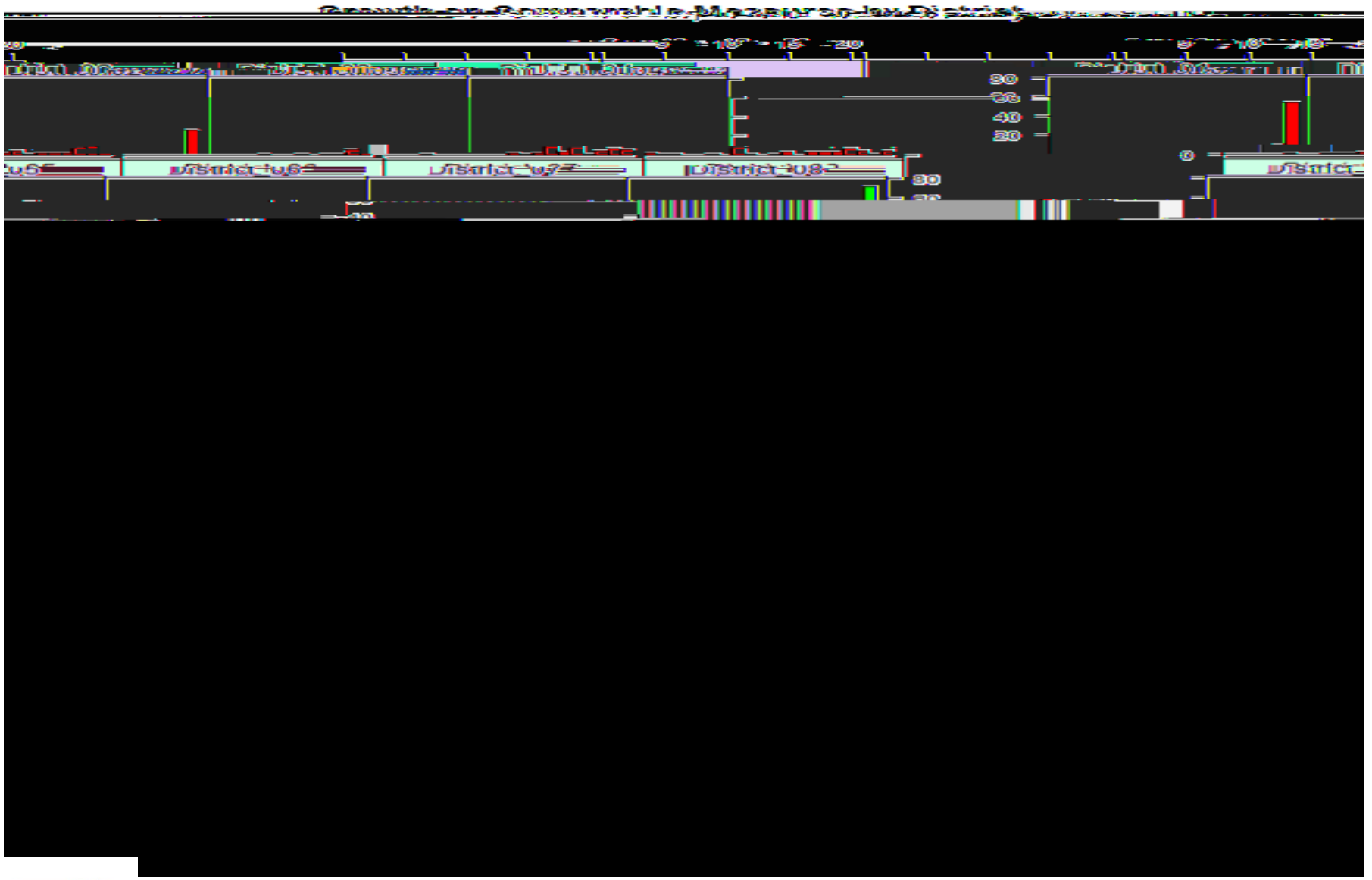

District_v6
District_v7
District_v8
80
60
40
20
0
80

Table Talk #3 (15 min)

Consider the following two scenarios:

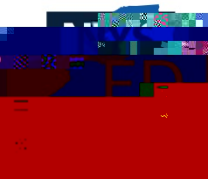
- Scenario 1: Student Growth remains an advisory measure, or are removed entirely and permanently from educator evaluation
  - If so: what (if anything) replace them as measures of student learning and educators' contributions to it (for example, SLOs)? Or, focus only on professional practice measures?
- Scenario 2: Expand Student Growth to other assessments (not just state tests)
  - This can be done, to some extent, with end-of-course exams (e.g., Hillsborough County)
  -

Guiding Questions

-

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



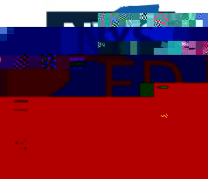
Appendix E

Evaluation Workgroup Session 3

NYSED Evaluation Workgroup

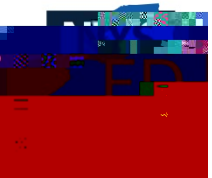
Webinar #1

Welcome!



Agenda

- Welcome and Introductions
- Agenda and Objectives
- Session #1 Debrief
- Current practices and policies regarding SLOs and growth models in NY
- History and Review of Student Growth Models and SLOs
- Break-out group discussion and share out of ideal student learning component of evaluation system
- Closing and Next Steps



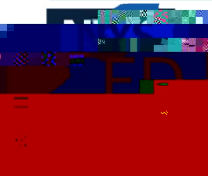
Objectives

- Understand the inputs provided during Session 1 and how they have informed the planning for the remaining two in person sessions
- Know the requirements for student performance measures specified in Education Law §3012

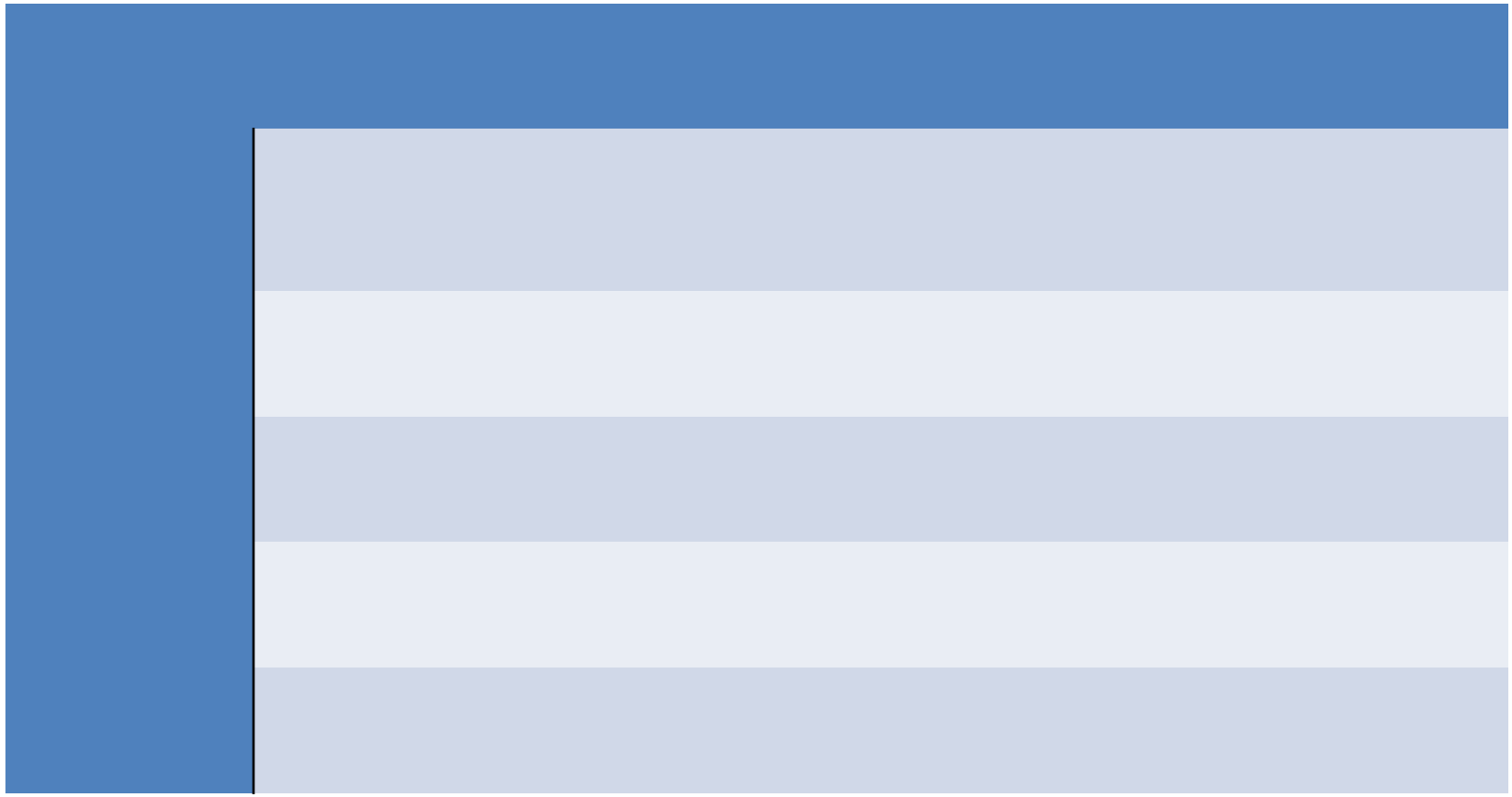
Session #1 Debrief

Purposes and Uses of an Ideal Educator Evaluation System

Educators seek an evaluation system that enhances teaching practice, especially through professional learning and growth opportunities, as well as via



Overview of Education Law §3012-d

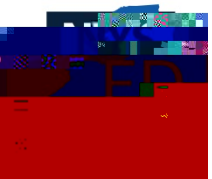


Education Law §3012-d

Student Performance Requirements

Required Measures

- Teachers of grades 4-8 ELA and math, principals of buildings covering these grade levels, and high school principals (all of grades 9-12) receive a State- provided growth score.
  - Statistical growth score calculated based on students' ELA and math State assessment results in the current year compared to similar students.
  - The term "similar students" in this context means not just students with the same academic history, but also students with the same demographic characteristics (i.e., English language

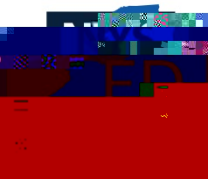


Education Law §3012-d

Student Performance Requirements

Required Measures

- All other teachers and principals have Student Learning Objectives (SLOs).
- An SLO is an academic growth goal set for an educator's students at the start of a course.
  - Represents the most important learning that is aligned to learning standards, as well as other school and district/BOCES priorities.
  - SLO growth targets must be specific and measurable, based on available prior student learning data. This baseline data may come from a variety of sources including pre-tests/pre-assessments and a student's prior academic history.
  - Educators' scores are based upon the degree to which the goals were attained, as evidenced by student academic performance at the end of the course.



Education Law §3012-d

Student Performance Requirements

Required Measures

- Where a course or grade level ends in a State-created or administered assessment, the Education Law requires that that assessment be used as the evidence for the SLO (e.g., grade 8 science, Regents courses, NYSAA, NYSESLAT).
- The required student performance measures must cover the majority of a teacher's students across all the courses/grades they teach.
- For principals, at least 30% of students enrolled in the building must be covered by the required measures.
- Some educators have a mix of State-provided growth scores and SLOs.

Education Law §3012-d

Student Performance Requirements

Required Measures

- The required student performance measures must cover the majority of a teacher's students across all the coursesu11

Education Law §3012-d

Student Performance Requirements

Optional Measures

- In addition to State-provided growth scores and/or SLOs, all school districts and BOCES also have the option to collectively bargain additional, “optional” student performance measures under the law.
- Under Education Law §3012-d, this second measure must be:
  - A second State-provided growth score based on a State-created or administered assessment; or
  - A growth score based on a State-approved assessment calculated using a State-approved growth model.
- Each measure assigns a score from 0-20, and the overall score corresponds to a rating of Highly Effective, Effective, Developing, or Ineffective (HEDI).

APPR Transition Period Requirements

- For the 2015–16 through 2018–19 school years, educators whose evaluations are to be based on the grades 3-8 ELA and math State tests and/or State-provided growth scores receive an “original” evaluation that includes these measures. This evaluation is for advisory purposes only.
- These educators also receive a “transition” evaluation that

Quick Review (cont’d)

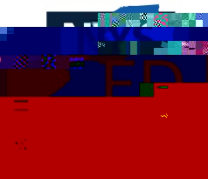
-

Poll

By taking into account factors that are outside of the teachers' control, student growth measures provide a fairer measure of teacher contributions to growth than attainment measures.

- Strongly Disagree
- Disagree
- No opinion
- Agree
- Strongly Agree

Comments, questions, wonderings or other thoughts?

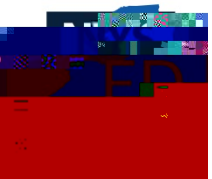


Student Growth Models : Concerns/Policy Issues (cont'd)

- Stakeholder concerns about state tests:
  - Year-to-year changes in state assessment systems and content standards

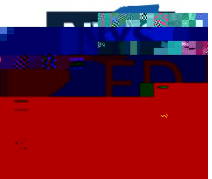
Student Growth Models: Concerns/Policy Issues (cont'd)

- High level of data complexity and effort needed to calculate statistical models accurately
- Reliability of value-added measures
- Causal attribution (i.e., Do models accurately capture teachers' true effect on student performance?)
- Use of school growth to indicate principal growth



Student Growth Models : Concerns/Policy Issues (cont'd)

- Models cannot explain why a particular teacher's students scored better than expected, so this measure is of limited use in a feedback-oriented system.
- Models are fundamentally different: a normative measure, whereas most other educator effectiveness measures are criterion-referenced.



Themes from Student Learning Objectives Implementation (Cont'd)

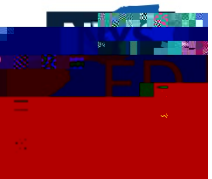
- Significant emerging challenges around:
  - Assessments:
    - Not enough high-quality assessments
    - Resistance to having lists of approved assessments
  - Growth Targets and Scoring:
    - Growth targets not necessarily informed by data
    - Potential incentive to set low targets
    - Scoring not consistent or comparable
  - Training and Support:
    - Not enough time for educators or evaluators to collaborate
    - Inadequate training on assessment development

Key Student Learning Objectives Decision Points for Policymakers

More Structure

Less Structure

Assessments/Evidence Sources
Growth Targets
Scoring Rubric

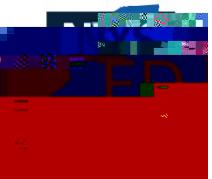


Decision Point 2:
Growth Targets and Scoring

- Tradeoffs of more/less structure:
  - State

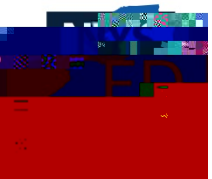
Additional Policy Considerations

- What kinds of resources and supports do districts and schools need to implement a high- quality SLO process?
  - Training: Initial and ongoing
  - Resources: Process guides, sample SLOs, etc.



Additional Policy Considerations (cont'd)

- How will scale- up of training take place?
  - Train-the-trainers models have benefits, but must ensure consistency and enough time
  - Concern about having only one trainer in a school/district who may also be a full-time teacher
  - How will longer-term capacity be built?



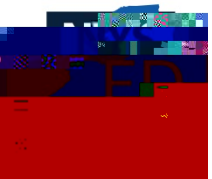
Key Questions (cont'd)

- Do not forget about data quality: Which SLO data will be entered into which platforms, and how will the data get integrated with other data sources to produce overall ratings?

Big Picture

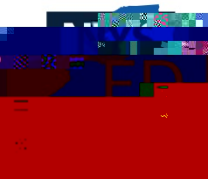
- Critical for all stakeholders to remember that all potential measures of educator effectiveness have tradeoffs
- There are challenges with student growth measures, but

Break-Out Group Discussion



Closing and Next Steps

- Thank you!!!
- Your inputs from today's webinar will be used to inform the planning for our next in person session
- January 10th.



Appendix F

NYSED Form H

Appendix G

NYSED sletemplate-3012-d