

# Integrating Information within a Balanced Assessment System

Nathan Dadey, Center for Assessment

Symposium on Learning-Focused Balanced Assessment Systems  
Presented at CCSSO's National Conference on Student Assessment  
June 28<sup>th</sup>, 2017





A simple premise:

*All* of data produced by a balanced assessment system should be use to characterize student learning.

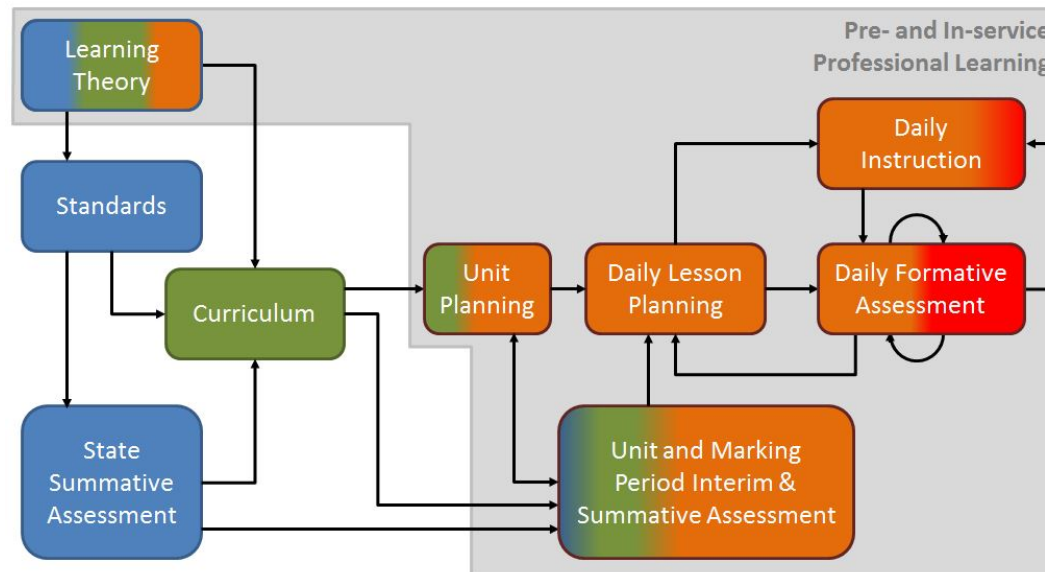


That is, if the *design* of a system is balanced, shouldn't the *analysis* be balanced too?



This sounds great, in theory, but is difficult in practice.

In large part, because of the power boundaries briefly touch upon in Joseph's presentation.





However, overcoming these types of barriers is important – otherwise we are leaving a lot of information on the table.



# Assessments Given in a Year

State. Large-Scale Standardized Accountability Assessment



Dist

How can we bridge these levels to better characterize student learning?

Classroom.



Quarter 1

Quarter 2

Quarter 3

Quarter 4



- Illustrate that examining data from two levels (state & district) can be powerful – providing insight that is, potentially, instructionally relevant,
- By investigating district data assessment data of sixth grade mathematics – data from three “modular” interims and a summative assessment.
- In the context of two uses outlined previously:
  1. Differentiating Instruction
  2. Auditing Grades



## Sixth Grade Mathematics

Summative

Interim 1



Interim 2



Interim 3



Q1

Q2

Q3

Q4

Start of  
Year

End of  
Year

- 3 interims with 30 items each & end-of-year summative (approx. 50 items)
- Interim items generally aligned to instruction in prior quarter (“modular” design)
- Approximately 5,000 students

# 1. Differentiating Instruction



## Intended Purposes and Uses of the District Assessment System

#	Priority	Description
1	mid	Audit District Assessment Results
2	mid	Audit Teacher/School-Assigned Marking Period Grades
3	high	Differentiate Instruction
4	mid	Evaluate Achievement for Traditional Grading
<p>Specifically, by <b>predicting proficiency classifications</b> on the summative assessment with interim assessment performance.</p>		
5	mid	Instructional Staff Training
10	low	Measure Educator/Institution Effect on Student Growth



# 1. Analysis & Results



- Used regression trees to find interactions of multiple variables that best predicted performance, then used those results to produce simple descriptive statistics.
- Findings
  - Students who scored 16 and above on the second interim (I2) were very likely to be proficient (91%)
  - Relationships for not passing were slightly more complex, e.g.,
    - Students with  $I2 < 16$  were likely to not be proficient (84%)
    - Students with  $I2 < 16$  &  $I3 < 16$  were very likely to not be proficient (0.92%)...

# 1. Results, Cont.



- So is the second interim important because of timing, content, or both?
  - Such information would need to be solicited empirically.
- However, the interims highly correlate, indicating that interim 1 or 3 could easily also be used to identify students based on a cutscore.



# 1. Caveats

- Prediction relationships may not:
  - transfer from cohort to cohort.
  - hold if educators act on them.
- Prediction  $\neq$  learning.



# 2. Auditing Grades

## Intended Purposes and Uses of the District Assessment System

#	Priority	Description
1	mid	Audit District Assessment Results
2	mid	Audit Teacher/School-Assigned Marking Period Grades
3	high	Differentiate Instruction

Specifically, by **comparing proficiency classifications** on the summative assessment with teacher grades.

9	mid	Instructional Unit Planning
10	low	Measure Educator/Institution Effect on Student Growth

# 2. Analysis & Results



- Columns sum to 100%.

Grade	Not Proficient	Proficient
A	6%	48%
B	23%	35%
C	30%	13%
D	22%	3%
F	18%	1%

# 2. Analysis & Results



- Columns sum to 100%.
- Implications for grading?

Grade	Not Proficient	Proficient
A	6%	48%
B	23%	35%
C	30%	13%
D	22%	3%
F	18%	1%

# 2. Caveats



- Begs further questions
  - Are there certain patterns of identifications that are undesirable? If so, are they grouped in schools, or perhaps can be explained by other student variables?



- Stop silo'ing data & start having conversations based on empirical examinations, hopefully guided by considerations of use like those in the tool put forth by Joseph.
- Wherever possible, posit hypotheses before conducting investigations.