Amphitheater's Teacher Performance Evaluation System (ATPES): How One District Approaches the Task of Teacher Evaluation Under the New Arizona Law 2013/2014



Roseanne Lopez, Ed.D.

#### Overview

- Overview of the teacher evaluation system
- ATPES Domains and Indicators
- Teacher Rating Scale
- Student Progress Metric: "OWA"
- School and Classroom "OWA" Reports
- Results/Observations
- Successes, Issues and Challenges

## Amphitheater Response A.R.S. 15-203

- Design Team work on principal, assistant principal and teacher evaluation systems
- Data analysis work/development on-going
- Approval by the Governing Board in June 2012
- "Year One" implementation 2012/2013
- "Year Two" implementation 2013/2014 with the addition of teacher labels and other changes in accordance with new state laws.

#### Important Background Facts

- Chose to modify our existing evaluation system rather than purchase or write a new one
- Amphitheater had a \$29 million dollar TIF Grant (2007-2012)
- Student progress data analysis model was developed beginning in 2007 utilizing familiar assessment information
- A Career Ladder program has been in place since the 1980's which requires Student Achievement Plans

#### Training of Evaluators

- Amphitheater ensures that all evaluators are qualified and Board approved.
- Required full-day training dates are established in the summer with three hour follow-up trainings once school has started. The emphasis of these workshops is inter-rater reliability.
- All evaluations are reviewed by the Associate Superintendent as they are completed; throughout the year, site administrators work on improving their skills in narrative writing and alignment with the indicators.

#### **ATPES Domains**

Domain 1:

Designing and Planning Instruction

Domain 2:

Assessing and Analyzing
Student Learning

Domain 3:

Creating and Sustaining the Learning Environment

Domain 4:

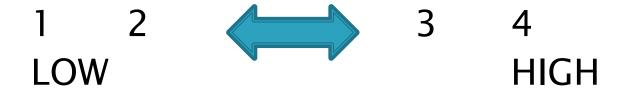
Implementing and Adjusting Instruction

Domain 5:

**Professional Responsibilities** 

### Rating Scale

ATPES now has a four point scale/rubric for teacher evaluation ratings.



A teacher at Level 1 may demonstrate a strategy for an indicator but it is incorrect or ineffective. A teacher at this level must improve his/her understanding and application of strategies for this indicator to correct or avoid negative effects on student progress.

Requires an Improvement Plan

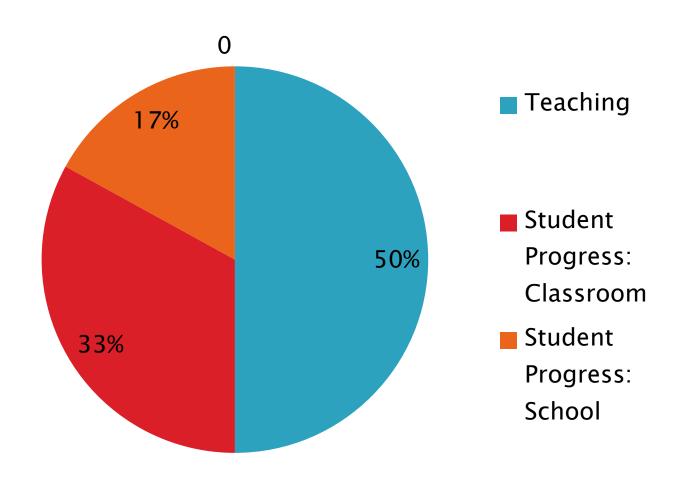
A teacher at Level 2 demonstrates strategies for an indicator but does not yet monitor its effectiveness on student learning. A teacher at Level 2 is progressing toward applying effective practices toward accomplishment of the indicator, but needs some assistance in mastering the skills necessary to positively affect student progress.

A teacher with several ratings of "2" should (at the very least) be on a Professional Development Plan. Recommendations for improvement from the administrator are required.

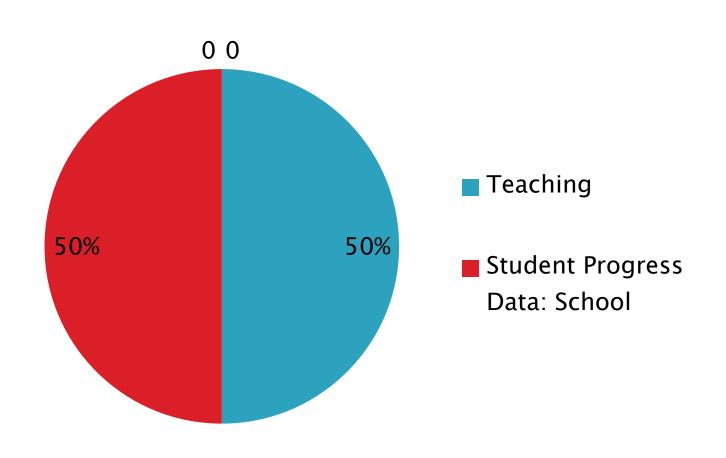
A teacher at Level 3 demonstrates appropriate best practices for an indicator, engages and guides students, differentiates and monitors the results. A teacher at this level is student-centered and focuses on the skills and strategies which result in increased student progress. The Level 3 teacher demonstrates mastery of this indicator.

A teacher at Level 4 exceeds the best practices of Level 3 through innovation, flexibility, and creativity for an indicator. The teacher at Level 4 is recognized as a leader in this area and could be called upon to model skills/competencies or assist in the development of this skill with colleagues.

## Group "A" Teachers



### Group "B" Teachers



#### School Level OWA

- High Schools
  - AZ Learns Label–School
  - AZ Learns Label–District
  - AIMS Math (Previous Three Years)
  - AIMS Reading (Previous Three Years)
  - NWEA MAP Reading (Previous Three Years as available)
  - NWEA MAP Math (Previous Three Years as available)
  - Advanced Placement Test Rate
  - Advanced Placement Pass Rate

#### School Level OWA

- Middle Schools
  - AZ Learns Label–School
  - AZ Learns Label–District
  - AIMS Math (Previous Three Years)
  - AIMS Reading (Previous Three Years)
  - NWEA MAP Reading (Previous Three Years)
  - NWEA MAP Math (Previous Three Years)

#### School Level OWA

- Elementary Schools
  - AZ Learns Label–School
  - AZ Learns Label–District
  - AIMS Math (Previous Three Years)
  - AIMS Reading (Previous Three Years)
  - NWEA MAP Reading (Previous Three Years)
  - NWEA MAP Math (Previous Three Years)
  - DIBELS (Previous Three Years)

#### Classroom Level OWA

- Teacher student progress OWA is available for "Group A" teachers only
- OWA at the classroom level is based on rubric scores for each data element over a three year time period
- Evaluators will enter the OWA at Indicator 2.5 and the instrument will automatically calculate the weighting

### Unpacking the OWA: AIMS Data

AIMS data will be utilized at the school level and at the individual teacher level (as deemed valid)

#### Analysis Description:

- Grades 3–8, 8 to 10
  - Observed Growth is the change from the spring scale score to the spring scale score the following year
  - Observed to Expected Growth is the average growth from one spring to the next spring of district students with similar characteristics. Students who have two data points that can be matched are included in the data set.

## Unpacking the OWA: Student Characteristics

- Starting Ability-Level
- Special Education
- English Language Learners
- Low Socioeconomic Status
- Mobility

...and any combinations of the above.

## Unpacking the OWA: NWEA MAP Data

- NWEA MAP data will be utilized in grades 2-9
- NWEA MAP data will be utilized at the school level and at the individual teacher level (as deemed valid)

#### Analysis Description:

- Observed Growth is the change in RIT score from the Fall to the Spring of the same year.
- Observed to Expected Growth is the average growth from the Fall to Spring of students with similar characteristics.
- Score is weighted by student count.

#### Unpacking the OWA: DIBELS Data

 DIBELS data will only be aggregated to the grade level, therefore, it will only be utilized for school level OWA

#### Analysis Description:

- Observed Growth is the difference in the composite overall score from one testing period to the next within the same year.
- Observed to Expected Growth is the average growth from one testing period to another of district students with similar characteristics.

#### Unpacking the OWA: School Labels

- ▶ The Arizona School Labels (A-F) are included to help keep a school-wide focus.
- OWA Rubric Scores are:
  - A=4
  - $\circ$  B=3
  - ∘ C=2
  - D=1
  - F=1
  - (There is no "0")

Note: There was a change in school labels recently. We will use the higher of the two labels if there is a difference.

## Unpacking the OWA: Advanced Placement

- Testing Rate: Number of students who took any AP test divided by the total enrollment
- Passing Rate: Number of tests passed at a score of 3 or higher divided by the number of tests attempted

#### Threshold:

- If the school tests 15% or more of the student enrollment, automatic score of 4 on the rubric.
- If 85% of testers pass at a score of 3 or higher, automatic score of 4 on the rubric.

# Observations from the Implementation

- A school with a high overall "OWA" rating can inflate an individual teacher's ratings
- A school with a low overall "OWA" rating can lower the score of a highly effective teacher

#### **Possible Corrections:**

Look at different percentage weights for classroom and school student progress results

# Observations from the Implementation

- Many teachers do not have valid and reliable student progress data for what they teach. They may be considered "Group B" for a very long time.
  - Example: The elementary physical education teacher who is responsible for teaching over 600 students per week in one 45-50 minute time slot per group.
  - Example: The secondary calculus teacher who has created exams and final exams that are exclusive to her courses.

# Observations from the Implementation

- There is less motivation to implement common end of course assessments in schools which may lead to "Group A" status for a teacher if the school where they teach has a high "OWA" score.
- Utilizing student growth rather than achievement is more equitable overall
- Use of three year trend data is more reliable, valid and fair for evaluation purposes

### Challenges

- We are being asked to develop a quantitative tool for a largely qualitative profession
- Developing and/or acquiring valid and reliable assessment measures for every curricular area
- Teacher training for understanding and growth
- Evaluator training for reliability

### Teacher Designation Labels

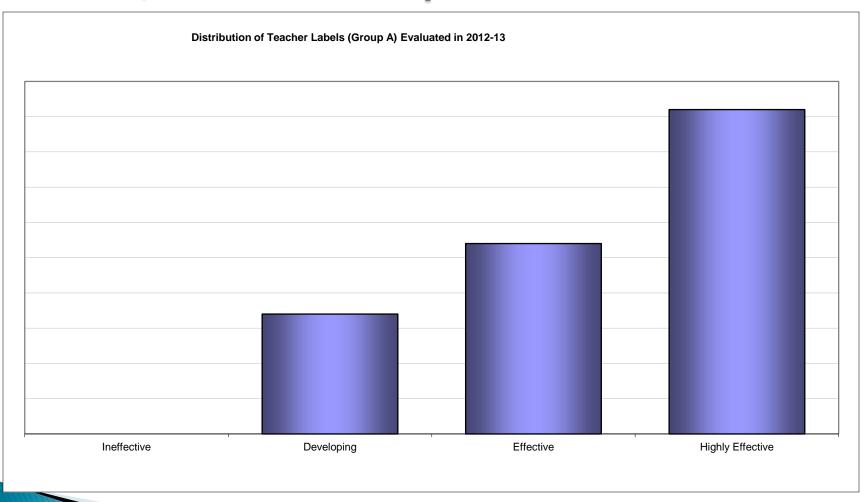
#### **GROUP A**

Criteria for effectiveness labels		
	Group A	
Highly effective	Total weighted rating=>119	
Effective	Total weighted rating between 113 and 118	
Developing	Total weighted rating between 100 and 112	
Ineffective	Total weighted rating < 100	

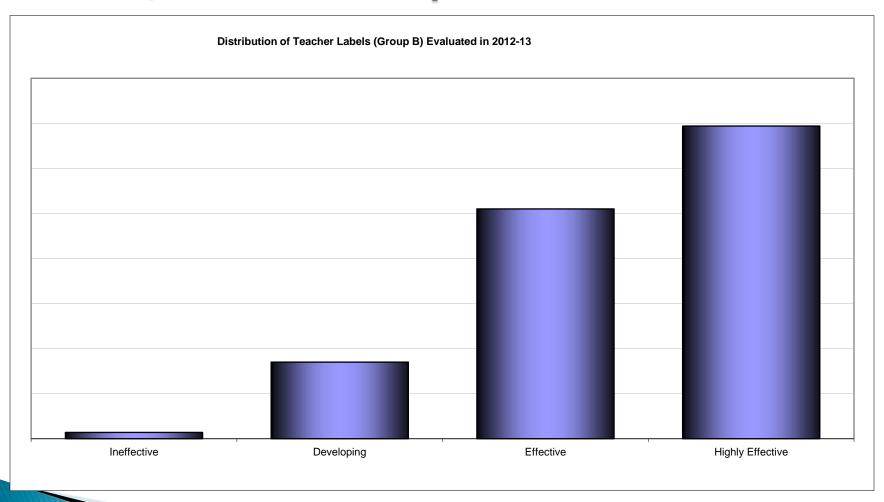
#### **GROUP B**

Criteria for effectiveness labels		
	Group B	
Highly effective	Total weighted rating=>106	
Effective	Total weighted rating between 94 and 105	
Developing	Total weighted rating between 78 and 93	
Ineffective	Total weighted rating < 78	

# Distribution of Teacher Labels 2012/2013 Group A



# Distribution of Teacher Labels 2012/2013 Group B

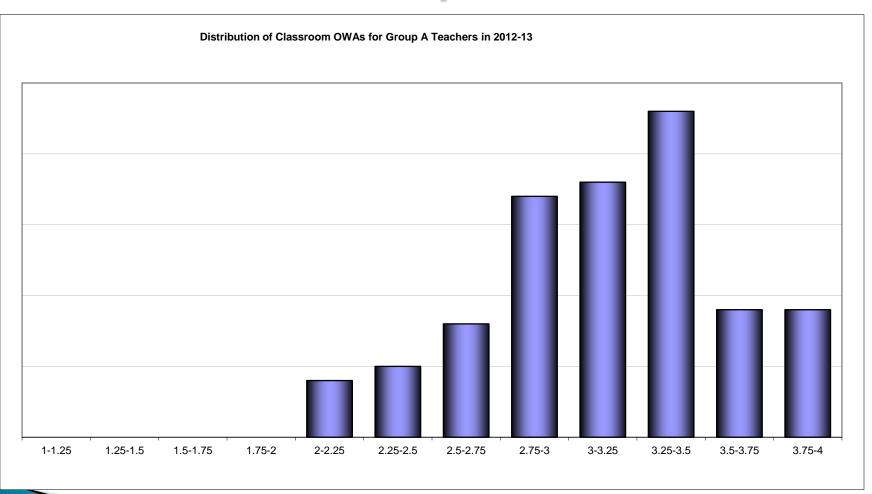


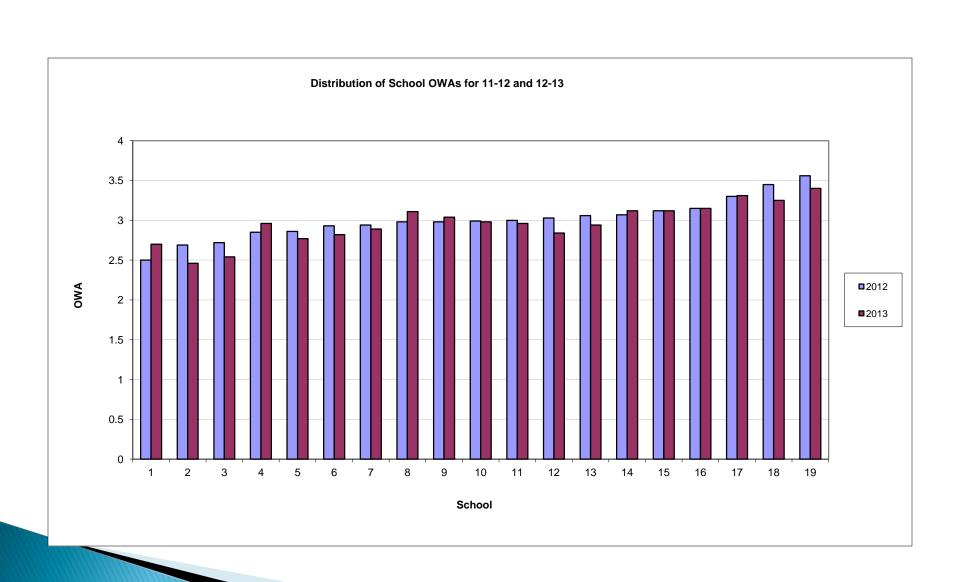
## Preliminary Results from <u>Year One</u> 2012–2013

Amphitheater "Group A" Teachers=10% Amphitheater "Group B" Teachers=90%

- Classroom student progress mean=3.14 (Group A)
- School student progress mean=2.98 (Group B)

# Distribution of Classroom OWA's 2012/2013 "Group A" Teachers





### Thank you!

- Roseanne Lopez, Ed.D.
  - rlopez@amphi.com
  - 696-5174

Monica Nelson mnelson@amphi.com Jay Midyett jmidyett@amphi.com



