



Championing Excellence!

**New Mexico's Teacher
and Principal Evaluation**

June 2013

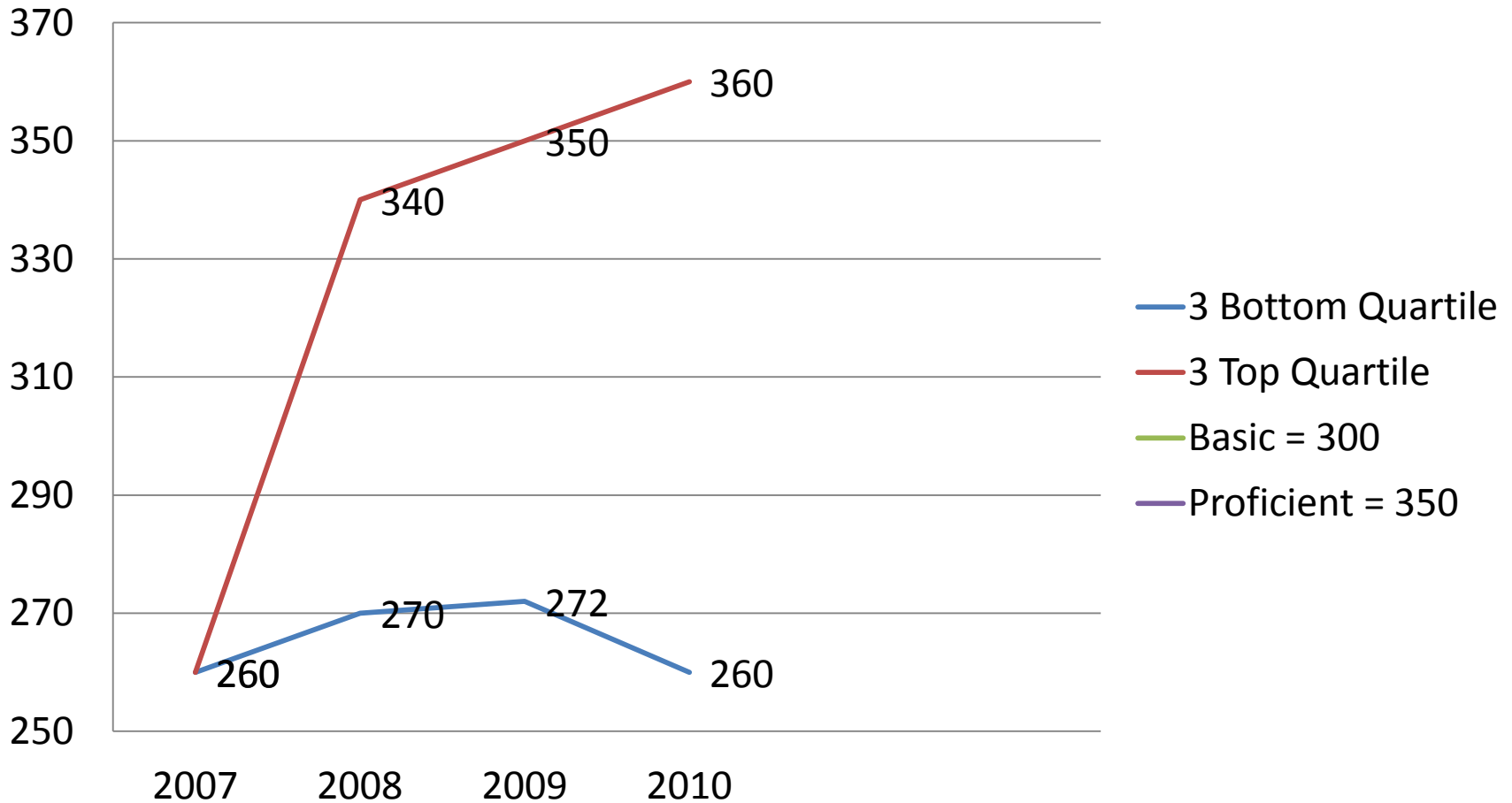
Teacher and Principal Evaluation

- The case for reform:
 - Currently over 99% of New Mexican teachers receive the highest rating of “meets competency” even though student achievement has not significantly increased.
 - Since 2003, New Mexican taxpayers have paid an additional \$400+ million in the Three Tier Licensure System for salary increases without corresponding improvement in student achievement.

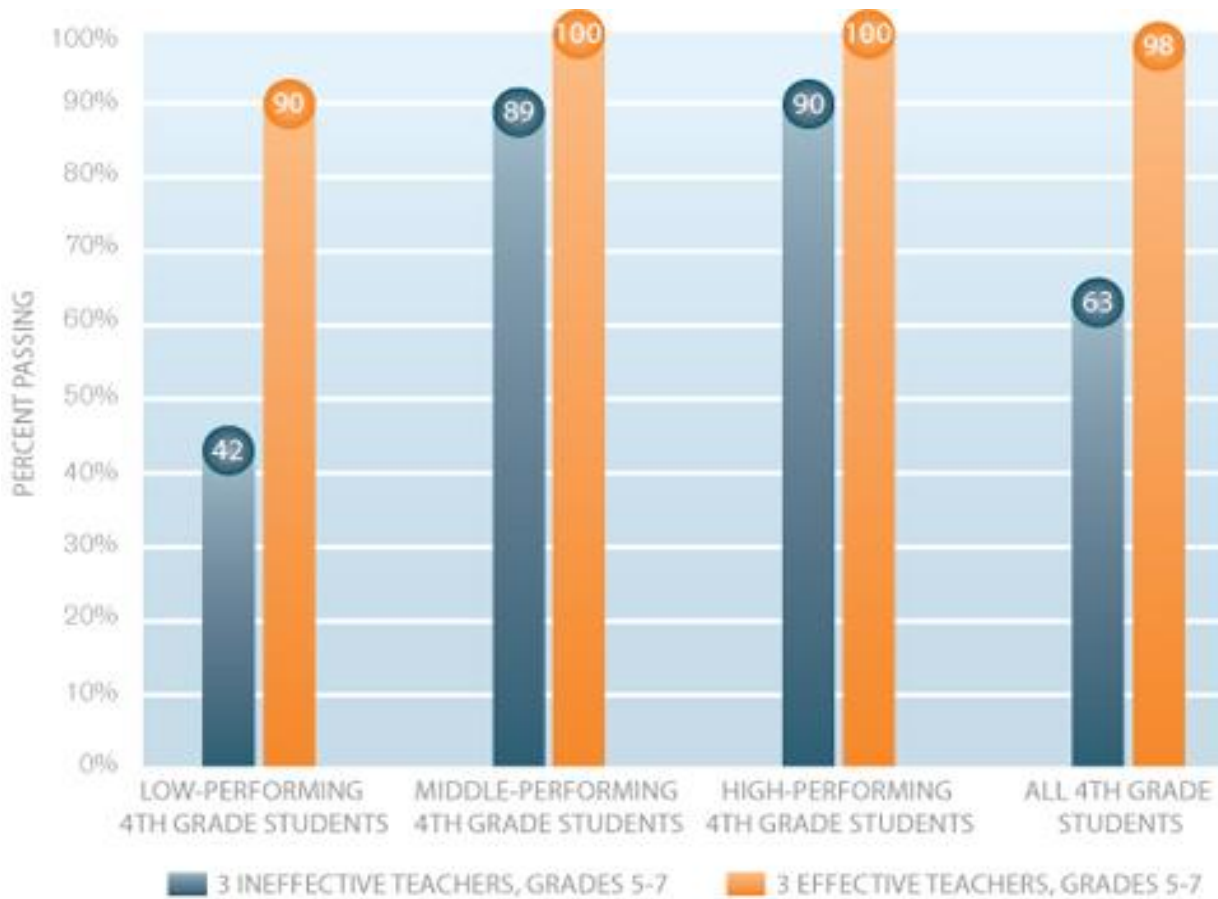
Three National Trends in Education

- The skill levels required for jobs in our communities are increasing.
- The educational gap between children of relatively affluent families and those of relatively poor families is widening.
- This generation of students is less likely to have more education than their parents.

Learning Denied



Effective Teachers Advancing Outcomes

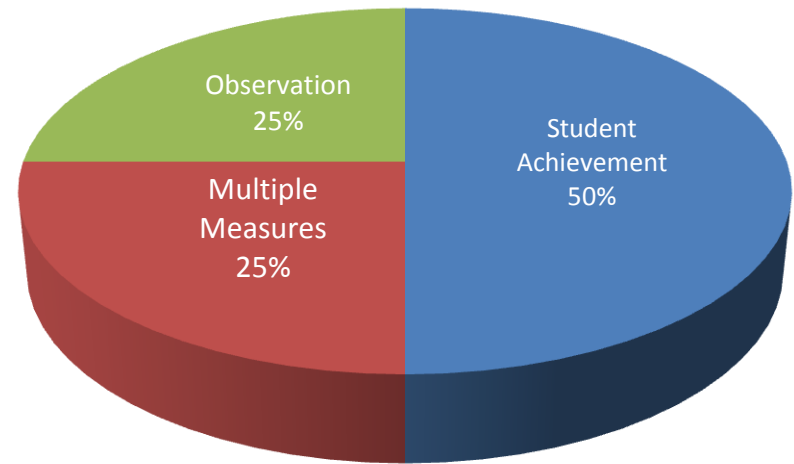


New Mexico Evaluation Framework

Teacher Evaluation

- 50% based on student achievement, of which:
 - 35% will be based on the SBA
 - 15% will be based on other measures of student achievement growth
- 25% based on observations
- 25% based on locally-adopted (PED approved) multiple measures

Teacher Evaluation Model (Tested Subjects)



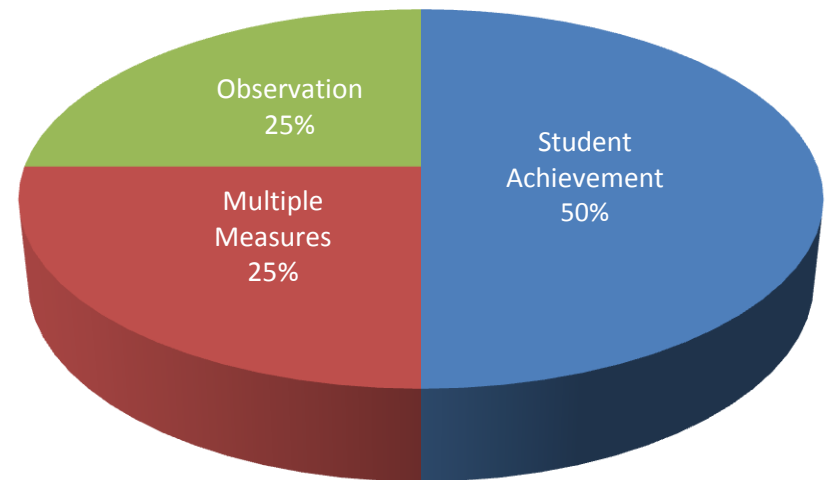
- Student Achievement
- Multiple Measures
- Observation

New Mexico Evaluation Framework

Teacher Evaluation

- 50% based on student achievement
 - Districts will submit relevant EOC, student growth measures, etc. to PED for approval
- 25% based on observations
- 25% based on locally-adopted (PED approved) multiple measures

Teacher Evaluation Model (Non-Tested Subjects)



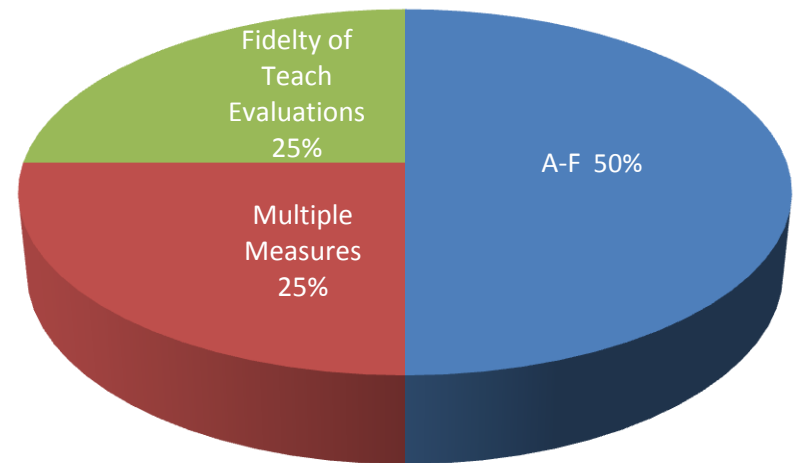
- Student Achievement
- Multiple Measures
- Observation

New Mexico Evaluation Framework

School Leader Evaluation

- 50% based on growth of a school's A–F School Grade
- 25% based on locally-adopted (PED approved) multiple measures
- 25% fidelity of teacher observations

School Leader Evaluation Model



- Growth in A–F School Grade
- Multiple Measures
- Fidelity of Teach Evaluations

Factors in Determining NMTEACH Evaluation Plans

- Prioritize student and school needs
- Focus improvement on locally determined priorities of achievement
- Improve overall school grade
- Capture differentiated performance among teachers
- Establish consistency and fairness
- Consider both short- and long-term implementation
- Determine feasibility of implementation

Graduated Considerations

- Kindergarten with no prior experience
 - Observation 75%
 - Multiple Measures 25%
 - No prior year achievement used

- All other grades

1st year teacher

Observations 50%
Multiple
Measures 50%

2nd year teacher

Achievement 25%
Observations 50%
Multiple
Measures 25%

3rd year teacher

Achievement 50%
Observations 25%
Multiple
Measures 25%

Group A Teachers

Tested Subjects and Grades:

- 3–5 All
- 6–8,
- 10–11 Language Arts/Math
- 6–7 and 9–11 Science
- Special Education-all grades (except services for students with severe and profound disabilities)

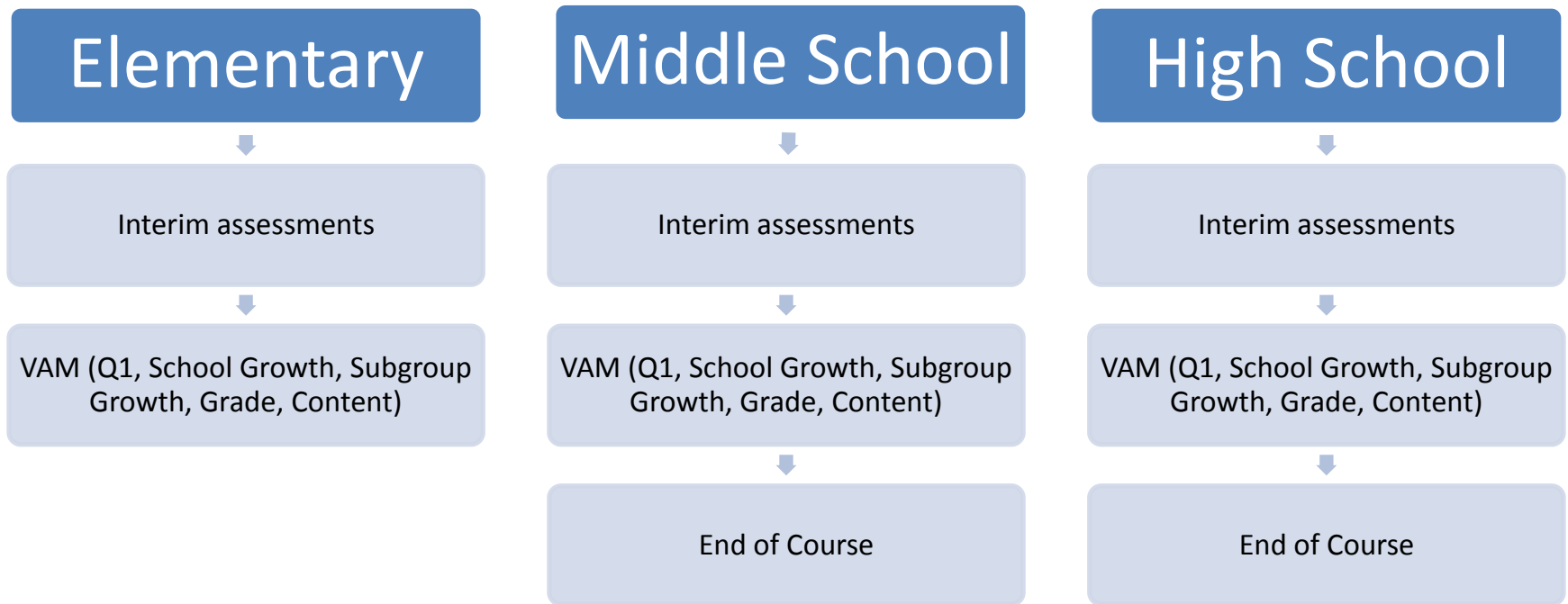
Standards Based Assessment (VAM)

- 35% based on growth
- 3 years of data (when possible)

Other Measures of Student Achievement

- 15% based on annual measures
- Interim assessments, End of Course Exams, Nationally Standardized assessments, “off the shelf” assessments

Determining Other Achievement Group A Possibilities



Student Achievement Example (100 pts)

VAM (70 pts)

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	20 and below	21–40	41–60	61–66	66–70

Algebra II-End of Course (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Class Average	10 and below	10–16	17–25	26–35	35–48

OR

Q1 Growth (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Letter Grade	F	D	C	B	A

Multiply Rubric Score by 6

Observations Example (50 pts)

Domain 2: Environment for Learning

Points	10 or less	11–14	15–19	20–23	24–25
--------	------------	-------	-------	-------	-------

Domain 3: Teaching for Learning

Points	10 or less	11–14	15–19	20–23	24–25
--------	------------	-------	-------	-------	-------

SUM

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	20 or less	21–29	30–38	39–46	47–50

Multiple Measures Example (50 pts)

Student Survey (25 pts)

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	21 or less	22–31	32–39	40–45	46–50

Total points divided by 2

OR

Teacher Attendance (25 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Days Missed	14+	11–13	6–10	3–5	0–2

Multiply rubric score by 5

Multiple Measures Example (*cont'd*)

(50 pts)

Domain 1: Planning and Preparation

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

Domain 4: Professionalism

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

SUM of Domains 1 and 4 DIVIDED by 2

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–14	15–19	20–22	23–25

Group B Teachers

- Non-tested Subjects and Grades:
- Elementary and Secondary non-core content teachers (i.e. Career Technical Education, Arts, etc.)
- 9 and 12 Language Arts/Math
- Secondary—Science/Social Studies

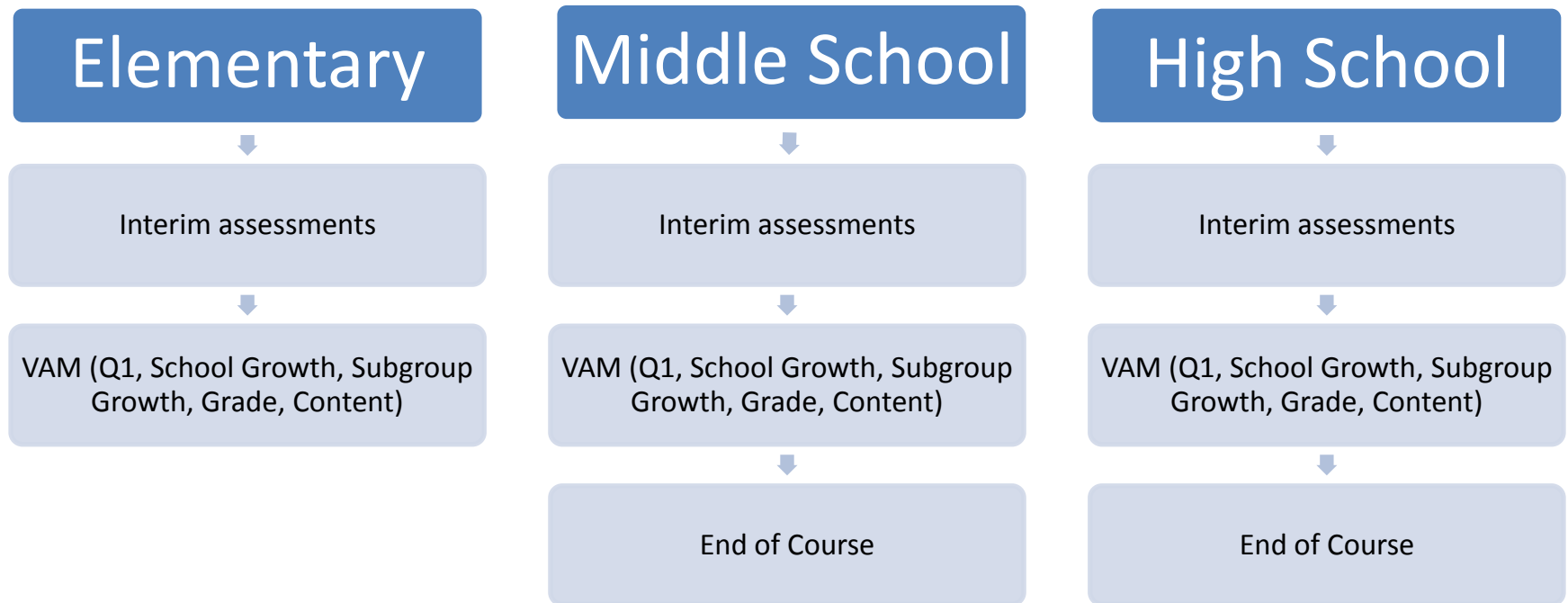
End of Course (or district selected achievement measure)

- 35% based on annual measures

Other measures of student achievement

- 15% based on annual measures
- Interim assessments, end of course exams, nationally standardized assessments, and “off the shelf” assessments

Determining Other Achievement — Group B Possibilities



Student Achievement Example (100 pts)

Algebra II-End of Course (70 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Class Average	10 and below	10–16	17–25	26–35	35–48
Multiply Rubric Score by 14					

Interim Assessment (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Annual Growth	< than 2 years	< than 1 year	1 year	1 year+	2 years
OR					

Q1 Growth (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Letter Grade	F	D	C	B	A
Multiply Rubric Score by 6					

Observations Example (50 pts)

Domain 2: Environment for Learning

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–15	16–19	20–22	23–25

Domain 3: Teaching for Learning

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–15	16–19	20–22	23–25

SUM

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	20 or less	21–30	31–38	39–46	47–50

Multiple Measures Example (50 pts)

Student Survey (25 pts)

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	21 or less	22–31	32–39	40–45	46–50

Total points divided by 2

OR

Teacher Attendance (25 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Days Missed	14+	11–13	6–10	3–5	0–2

Multiply rubric score by 5

Multiple Measures Example (*cont'd*)

(50 pts)

Domain 1: Planning and Preparation

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

Domain 4: Professionalism

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

SUM of Domains 1 and 4 DIVIDED by 2

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–14	15–19	20–22	23–25

Group C Teachers

Non-tested Grades: K–2

Interim Assessments

- 35% based on annual measures

Other measures of student achievement

- 15% based on annual measures
- Interim assessments, nationally standardized assessments, and “off the shelf” assessments

Determining Other Achievement — Group B Possibilities

Elementary



Interim assessments



VAM (Q1, School Growth,
Subgroup Growth, Grade, Content)

Student Achievement Example (100 pts)

Interim Assessment (70 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Annual Growth	< than 2 years	< than 1 year	1 year	1 year +	2 years
Multiply Rubric Score by 14					

End of Year (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Points	50	60	70	80	90

OR

Q1 Growth (30 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Letter Grade	F	D	C	B	A
Multiply Rubric Score by 6					

Observations Example

(50 pts)

Domain 2: Environment for Learning

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–15	16–19	20–22	23–25

Domain 3: Teaching for Learning

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–15	16–19	20–22	23–25

SUM

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	20 or less	21–30	31–38	39–46	47–50

Multiple Measures Example (cont'd)

(50 pts)

Student Survey (25 pts)

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	21 or less	22–31	32–39	40–45	46–50

Total points divided by 2

OR

Teacher Attendance (25 pts)

	Ineffective (1)	Minimally Effective (2)	Effective (3)	Highly Effective (4)	Exemplary (5)
Days Missed	14+	11–13	6–10	3–5	0–2

Multiply rubric score by 5

Multiple Measures Example (cont'd)

(50 pts)

Domain 1: Planning and Preparation

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

Domain 4: Professionalism

Points	10 or less	11–14	15–19	20–22	23–25
--------	------------	-------	-------	-------	-------

SUM of Domains 1 and 4 DIVIDED by 2

	Ineffective	Minimally Effective	Effective	Highly Effective	Exemplary
Points	10 or less	11–14	15–19	20–22	23–25

MET Project

- Observations as part of a multiple measure system
 - Define expectations for teachers
 - Ensure observer accuracy
 - Ensure reliability of results
 - Determine alignment with outcomes

MET Project

- Two criteria for conducting formal observations
 - Reliability
 - Results reflect consistent aspects of a teacher’s practice
 - Do **not** reflect the idiosyncrasies of a particular observer, group of students, or lesson
 - Validity
 - The extent to which observation results are related to student outcomes

Lessons Learned

- All five instruments were positively associated with student achievement gains
 - Teachers who demonstrated effective practices in the classroom also had greater student gains in standardized assessments
- Reliability characterizing a teacher's practice required averaging scores over multiple observations
 - Multiple raters performing multiple observations creates greater reliability than a single rating by one observer

Training and Support

- In June and July, the PED will provide 9 regional 2–day institutes on the NMTEACH observation protocol.

June 3–4	Albuquerque
June 10–11	Farmington
June 17–18	Las Vegas
June 19–20	Portales
June 26–27	Hobbs
July 8–9	Silver City
July 10–11	Las Cruces
July 22–23	Santa Fe
July 24–25	Albuquerque (NMCSA)

Training and Support

- In August, the PED will provide guidance updates via webinar
- Beginning in September, the PED will provide four, one day regional training sessions per month

Questions

Matt Montano

Director of Educator Quality

matthew.montano1@state.nm.us

505-827-6581