



Framework for a Principal Evaluation System

Andrew Porter – University of Pennsylvania

Joseph Murphy – Vanderbilt University

Ellen Goldring – Vanderbilt University

Steven Elliott – Arizona State University

Step 1

Decide on the dimensions of information that are relevant to the task of evaluating principals for the purposes you have in mind. We strongly recommend that the first two dimensions be (1) principal leadership behaviors aligned with the ISLLC standards (e.g. VAL-ED) and (2) school outputs.

Step 2

Select reliable and valid measures of each dimension of the evaluation (e.g. value added to student achievement and principal leadership behaviors). A principal evaluation system should have at least one measure per dimension, but there can be multiple measures for a dimension or measures that cut across dimensions. What follows are some examples from four districts. As you can see, their measures tap into principal leadership behaviors and school outputs.

Measures	Districts			
	District 1	District 2	District 3	District 4
Dimension 1: Leadership Behaviors				
Principal Goals	X	X	X	X
VAL-ED 360° Assessment	X	X	X	X
Dimension 2: School Outputs				
Value Added to Student Achievement	X	X	X	X
Customer Satisfaction	X			
Graduation and/or retention	X		X	

Step 3

Establish the value of the measures of the dimensions used to assess the principal by assigning weights to them. Typical weighting schemes assign a portion of 100% of the total score to each measure based on a consensus view of the relative importance of each measure. Two illustrative examples follow:

Example #1:

Organizational goal attainment	15%
Student achievement	40%
360° assessment (VAL ED)	30%
<u>Professional developmental goal attainment</u>	<u>15%</u>
Total	100%

Example #2:

360° assessment (VAL ED)	50%
<u>Student achievement</u>	<u>50%</u>
Total	100%

Step 4

Define performance levels that characterize the entire range of performances within the evaluation framework. We recommend using four levels (e.g., below basic, basic, proficient, distinguished) for characterizing the range of performances where Below Basic represents the lowest level of performance and Distinguished represents the highest level of performance. In most cases, each performance level is described by a narrative rubric and then operationalized by scores/ratings for the total score of the set of measures.

Each performance level should be associated with a band of scores or set of ratings that are used to objectively assign a principal's performance on a given measure to a specific performance level. We illustrate some score bands that could reasonably define the boundaries of each performance level.

Example #1: Student achievement

Below basic	Less than +.9 years growth
Basic	+ .90 to +1.19 years growth
Proficient	+1.20 to 1.29 years growth
Distinguished	More than +1.29years growth

Example #2: 360° assessment (VAL ED)

Below basic	1:00 to 3:28 Mean effectiveness rating
Basic	3:29 to 3:59 Mean effectiveness rating
Proficient	3:60 to 3:99 Mean effectiveness rating
Distinguished	4:00 to 5:00 Mean effectiveness rating

The performance levels for each measure serve as a common scale for comparing results for the various measures. That is, with four performance levels for each measure, one can state that performances at the Below Basic level for both Student Achievement and the 360° assessment (VAL ED) earn a principal a 1, whereas when the performances fall within the Distinguished level for two such indicators the principal would earn a 4.

Step 5

Once all measures for a principal are placed on the performance level continuum where they each have a value of 1 through 4, they can be weighted according to the established values and a Total Evaluation Score can be determined. The example below provides a weighted performance level score for each measure and then the Total Evaluation Score.

Example: Measures with Weightings

Organizational goal attainment (15%)	Basic (2)	30/60
Student achievement (40%)	Basic (2)	80/160
360° assessment (VAL ED) (30%)	Proficient (3)	90/120
Professional development goal attainment (15%)	Proficient (3)	45/60
Total Evaluation Score		<u>245/400 (or 61%)</u>

The interpretation of this Total Evaluation Score can be accomplished in several ways. An example that logically extends the practices associated with the four level performance continuum is illustrated below:

<i>Example:</i>	Below basic	0%-50%
	Basic	51%-65%
	Proficient	66%-84%
	Distinguished	85%-100%

Step 6

Evaluations have consequences for users and the person being evaluated. Key issues to consider when evaluating principals include:

- When using the VAL-ED as one of the measures for principal evaluation with high stakes involved (e.g. promotion, tenure, merit pay), should the principal's own evaluation and the supervisors count as well, or should the assessment rely on just teacher responses?
- When deciding on performance levels, should the performance levels be set on each measure, or should a composite of measures be built and the performance levels set on the composite (i.e. the dimension)?
- When making cut scores for decisions about merit increases or bonuses, should the performance levels be set against a particular standard, such as proficient, or should they be set against a norm-referenced standard, such as the top ten percent being evaluated?
- Should the principal's length of time in the school be taken into account when deciding how much weight to give to value added to student achievement? Perhaps the VAL-ED should be weighted as much as 75% for principals new to a building and as little as 25% for principals that have been in the building six years or more.
- The VAL-ED has sufficient psychometric quality to be used for high stakes decisions (e.g. merit pay) but also can serve formative/diagnostic purposes.

The figure below provides a basic model for assessment of leadership in K-12 education. The shaded box represents principal behaviors as assessed by VAL-ED. They translate into school performance which translates into school outputs. The box at the bottom represents characteristics of the context that might be considered when interpreting performance on the VAL-ED and student success.

