Key Components in Principal Evaluation Systems Design

Matthew Clifford, Ph.D.
Senior Researcher

Supporting State Efforts to Implement Comprehensive Teacher and Leader Evaluation Systems
January 26, 2012
Why focus on principal evaluation system improvement?

Improvement is a discipline, a practice that requires focus, knowledge, persistence, and consistency over time.

- Richard Elmore

90,000 public school principals
98,706 public schools
3 million teachers
55 million PK-12 public school students

U.S. Census, 2011; Battle and Gruber, 2010; Gates et al., 2002
Why focus on principal evaluation system improvement?

- Creating and sustaining an ambitious vision and mission
- Engaging with teachers and data to improve instructional quality and access.
- Efficiently managing resources.
- Creating safe learning environments for students and staff.
- Developing strong and respectful community relationships.
- Acting in a professional and ethical manner.

35 states have adopted or adapted ISLLC standards.

NCATE adopted ISLLC standards for preparation program certification.

CCSSO, 2008; Marzano, Walters, & McNulty, 2005; Strong, Richard & Catano, 2008
Why focus on principal evaluation system improvement?

**Federal initiatives**
- Race to the Top (RTTT)
- Teacher Incentive Fund (TIF)
- School Improvement Grants (SIG)

**State policies**
- 31 states recently passed principal evaluation improvement policies

As written, policies emphasize accountability and learning. Thus, providing assessments of principal quality and feedback for improvement.

Almanzan and Kearny, 2011; Brown-Sims, 2011; Clifford and Kimball, in process
Why focus on principal evaluation system improvement?

- Principals report having **few sources of feedback**.
- Principals view evaluation as having **little impact** on their sense of accountability or practice.
- Performance assessments are **inconsistently administered**.
- Performance assessments are not **often aligned** with existing professional standards, and lack psychometric research.
- Principal performance assessment is not implemented in ways that maximize rating consistency, validity, and impact.

Why Evaluation Design Matters and How We Can Help

- **Validity**: Assures standards alignment and systems feasibility, both threats to evaluation validity.

- **Fidelity**: Identifies resources and methods for systematic implementation across contexts, which contributes to validity and reliability.

- **Trustworthiness**: Builds political support for improved evaluation.

- **Compliance**: Federal initiatives system design processes to:
  - Involve constituents
  - Clearly and broadly communicate systems requirements
  - Assure data quality
  - Align evaluation with principal learning systems
  - Reflect on system performance.
Why Evaluation Design Matters and How We Can Help

• Focus on evaluation for learning.

• Provide case examples of systems and system design.

• Present guidelines and criteria for strong system design.

• Help facilitate processes and supply useful research and other resources.

• Define differences between principal and teacher evaluation (and highlight ways that systems can be mutually supportive and coherent).

• Connect state/district designers to researchers and advocates.

• Emphasize importance of systems testing and quality.
Practical Guide to Designing Comprehensive Principal Evaluation Systems

What the Guide does:
- Facilitates informed decision-making
- Summarizes national policy contexts
- Synthesizes research on performance evaluation systems design and leadership effectiveness
- Focuses on key components and questions that drive evaluation design
- Provides links to TQ Center resources (e.g., Guide to Evaluation Practices), state/district case examples, definitions of terms
- Will be accompanied by an online, interactive tool

How the Guide was written:
- Informed by state and district design processes
- Research review
- Written and reviewed by evaluation systems designers, national professional association staff, and researchers
Key Components for Principal Evaluation Systems Design

Component 1: Specifying Evaluation System Goals

Component 2: Making a Strategic Communication Plan

Component 3: Selecting Measures

Component 4: Determining the Structure of the Evaluation System

Component 5: Selecting and Training Evaluators

Component 6: Ensuring Data Integrity and Transparency

Component 7: Using Principal Evaluation Results

Component 8: Evaluating the System

For criteria on strong evaluation systems, see Committee for Standards in Educational Evaluation, 2010; Leon, Davis, Sanders, Kearney, and Thomas, 2011; NAESP and NASSP, in process
Component Example: Specifying System Goals

Overarching Design Questions

1. What are the goals and purposes of the evaluation system?
2. Are the goals explicit, well-defined and clearly articulated for stakeholders?
3. Have the evaluation goals been aligned to the state strategic plan, the principal evaluation systems design communications plan, principal preparation and professional development initiatives, and pertinent school improvement initiatives?
4. What is the definition of an effective principal, and how does it align with the definition of an effective teacher?
5. What standards will drive evaluation design?

Weighing Goal Choices
- Improvement of practice?
- Decisions about competency?
- Articulating state/district priorities?
Component Example: Selecting Measures

Overarching Design Questions

1. Have measures been selected in light of key criteria (e.g., alignment to system purposes, strength of measures, use across contexts)?
2. Will student growth in tested subjects and other outcomes measures be factored into principal evaluation, and if so, how will they factor into principal evaluation?
3. Will student growth in non-tested subjects be included in principal evaluation, and if so, what weight will they be given?
4. To what degree will principal practice measures (e.g., observations, school climate surveys, 360-degree evaluations) be included?

Embedded Questions
What is a “technically sound” measure?
What are commonly used practice and outcomes measures?
How, if at all, should the evaluation system be differentiated?
What measures should be used, and when, to evaluate?
How evidence be weighted?
<table>
<thead>
<tr>
<th>Domain</th>
<th>Practice Measurement Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Creating a mission and vision</td>
<td>School climate survey&lt;br&gt;360-degree assessment&lt;br&gt;Evidence of SIP progress</td>
</tr>
<tr>
<td>Improving instructional quality</td>
<td>Observations of teacher evaluation, data review&lt;br&gt;Instructional quality measures&lt;br&gt;Evidence of SIP progress&lt;br&gt;360-degree assessments&lt;br&gt;School climate survey</td>
</tr>
<tr>
<td>Efficiently managing resources</td>
<td>Fiscal review&lt;br&gt;Safety and compliance record</td>
</tr>
<tr>
<td>Creating safe learning environments</td>
<td>Student survey&lt;br&gt;School climate survey&lt;br&gt;Evidence of SIP progress&lt;br&gt;Parent/community survey</td>
</tr>
<tr>
<td>Developing strong community relationships</td>
<td>Evidence of SIP progress&lt;br&gt;Parent/community survey</td>
</tr>
<tr>
<td>Acting in a professional and ethical manner</td>
<td>360-degree assessments&lt;br&gt;Portfolio review</td>
</tr>
</tbody>
</table>

Examples from Iowa, North Carolina, Tennessee, Delaware, New York City, Hillsborough, Pittsburgh, Round Rock
## Outcomes Measures Alignment

<table>
<thead>
<tr>
<th>Domain</th>
<th>Outcomes Measurement Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>School Conditions</td>
<td>School incident reports</td>
</tr>
<tr>
<td></td>
<td>School climate surveys</td>
</tr>
<tr>
<td></td>
<td>Working conditions surveys</td>
</tr>
<tr>
<td></td>
<td>Parent/community survey</td>
</tr>
<tr>
<td>Teaching Quality</td>
<td>Teacher certification and personnel data</td>
</tr>
<tr>
<td></td>
<td>Prof Development Plan data</td>
</tr>
<tr>
<td>Instructional Quality</td>
<td>Learning quality measures</td>
</tr>
<tr>
<td></td>
<td>Teacher evaluation results</td>
</tr>
<tr>
<td></td>
<td>Teacher surveys</td>
</tr>
<tr>
<td>Student Learning</td>
<td>VAM</td>
</tr>
<tr>
<td></td>
<td>Student growth measures</td>
</tr>
<tr>
<td></td>
<td>State school data</td>
</tr>
</tbody>
</table>

Examples from Iowa, North Carolina, Tennessee, Delaware, New York City, Hillsborough, Pittsburgh, Round Rock
Example of Weighted Evidence
Hillsborough County School District (FL)

Weighting evidence and reporting results

- 40% value added scores
- 30% 360-degree evaluation
- 30% local measures

- 10% - Student Attendance & Discipline
- 10% School Operations (finance and facilities)
- 5% teacher retention
- 5% teacher evaluation results
Matthew Clifford
P: 630-689-8017
E-Mail: mclifford@air.org

20 North Wacker, Suite 1231
Chicago, IL 60606
Website: www.tqsource.org