Special Education Teacher Evaluation: Issues and Answers

Tony Bagshaw
Battelle for Kids

&

Lynn Holdheide
Vanderbilt University

National Association of State Directors of Special Education Professional Development Series

Copyright © 2010 National Comprehensive Center for Teacher Quality. All rights reserved.
Learning Targets

Seeks to build the capacity of participants to

• Articulate the challenges identified with evaluating special education teachers through value-added and other measures of teacher evaluation.

• Actively participate in the creation or redesign of teacher evaluation models that support the development of strong, valid and reliable teacher evaluation policies and practices that recognize and promote the unique contribution of special education teachers.
A Reflection over the Last 100 Years

- **1900**: Teacher evaluations mostly @ personal qualities (grooming, articulation, confidence, etc.)
- **1950’s**: Teacher evaluations mimic industry appraisals (checklists, inventories, etc.), getting more formalized in nature.
- **Mid-1960’s**: Coleman et al. (1966) Unflattering study on the effects of schools
- **1970’s**: Madeline Hunter influenced Teacher Evaluations
- **Late-1980’s**: More studies cast doubt (100 district study)
- **1990’s**: Danielson’s Framework (1996)
- **1996**: 100 district study replicated – no change/improvement
- **Mid-1990’s - now**: Teacher Effectiveness studies start emerging, showing profound impact of teachers on student learning
- **Now**: CLASH! of subpar (lousy) teacher evaluation/improvement systems WITH Knowledge of Teachers Quality Importance!

www.BattelleforKids.org
Concerns and Attention Mount

Policy Papers & Foundation Efforts

- Education Sector’s (2008) *Rush to Judgment*
- New Teacher Project’s (2009) *Widget Effect*
- Center for American Progress (2009) “So Long, Lake Wobegone”
- Gates Foundation funds 4-site effort to develop teacher effectiveness measures

Policy

- State Fiscal Stabilization Funds (4 assurances, Great teachers/leaders)
  - LEA’s might use SFSF money to “[establish] fair and reliable evaluation systems that provide feedback, help educators improve, and ensure that poor performers are dismissed”
- Race to the Top (4 assurances) went further ..
  - (D)(2) Improving teacher and principal effectiveness based on performance (58 points)
- White House ESEA Reauthorization Recommendations
Value-added research shows that teachers vary greatly in their contributions to student achievement and teacher effectiveness is the most influential school-based factor in student achievement (Rivkin, Hanushek, & Kain, 2002; Sanders & Horn, 1998; Sanders & Rivers, 1996).
Summary of Teacher Evaluation Problems

- All teachers are rated as good or great. Because of this...
  - Excellent performance goes unnoticed
  - Typical goes without support to improve further
  - Chronically low performing goes unaddressed

- Results of Teacher Evaluation have little/no impact on HR decisions
  - Retention, promotion, placement, compensation, professional development, tenure, etc.

- Result: Schools’ #1 factor for making a difference with students is treated indifferently, so education’s effect on student outcomes is likely compromised (heavily?)
Guidance for Improving Teacher Evaluation

- Professional Evaluation Standards
- State Law / Tennessee among others
- Evaluation Purposes / Summative vs. Formative
- SEA efforts and LEA efforts abound and are on the fast track
Impact on Special Educators

- Special educators included in the accountability mandates
- Inclusion in performance-based compensation systems
- Need to identify the special challenges in evaluating special educators
- Determine where systems fit and/or need to be differentiated
Persistent Challenges for Special Educators

- Persistent achievement gap for students with disabilities
- Issues with teacher retention, recruitment, and attrition (McLesky & Billingsley, 2008)
- Special education positions are left vacant or filled with uncertified personnel (Billingsley, Fall, & Williams, 2006; Boe & Cook, 2006)
- A limited use of evidenced-based practices (Reschly, Holdheide, Smart, & Oliver, 2007; Walsh, Glaser, & Wilcox, 2006)
Contextual Challenges for Special Educators: Opportunity for effectiveness?

- Formidable content & pedagogical demands
- Insufficient time on instruction: 40% (Vannest & Hagan-Burke, 2010)
- Diverse & increasing caseloads (Carlson et al., 2002; McLeskey et al., 2004)
- Special education teachers often have little input into co-teaching and collaborative instruction & a general lack of clarity about roles (Gehrke & Murri, 2006)

Slide courtesy of Bonnie Billingsley, Virginia Tech
The Purpose


- Identify the specific challenges in evaluating this population of teachers.
- Determine the current status of state policy and practice.
- Identify promising evaluation practices and instruments.
- Provide guidance and policy recommendations to districts and states.
Review of policy/literature
Survey inquiry
Series of interviews with state- and district-level practitioners and researchers
Data collection period: December 2009–April 2010

- Designed in collaboration with Council for Exceptional Children (CEC) and national experts
  - State and local survey
  - Respondent pool: state and local directors (identified within CEC’s Council of Administrators of special education listserve)

1,143 total respondents
Table 3. District’s Current Evaluation System

<table>
<thead>
<tr>
<th>Statement</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our district uses the state’s recommended evaluation system.</td>
<td>200</td>
<td>22.1</td>
</tr>
<tr>
<td>Our district uses a slightly modified version of the state’s evaluation system.</td>
<td>129</td>
<td>14.2</td>
</tr>
<tr>
<td>Our district has developed our own teacher evaluation system.</td>
<td>496</td>
<td>54.7</td>
</tr>
<tr>
<td>None of these efforts describe our state efforts. If so, please describe:</td>
<td>46</td>
<td>5.1</td>
</tr>
<tr>
<td>Don’t know</td>
<td>35</td>
<td>3.9</td>
</tr>
</tbody>
</table>

*Note: Responses were included for only those local special education administrators who indicated familiarity with the district teacher evaluation system.*
Among the local administrators, 81% reported that contractual agreement prevented modification in the evaluation process.
Meeting the needs of “diverse” learners may not attend to the following:

- Special skills (individualized education program [IEP] facilitation, collaboration, secondary transition, social and behavioral interventions, compliance with legal mandates)

- Evidence-based instructional methods (direct/explicit instruction, scientifically based reading instruction, learning strategy instruction)
Opinions Regarding Special Education Teacher Evaluation

- Special educator use of evidence-based strategies should be a component of the evaluation process: 92%
- Special educators should be evaluated using the same evaluation process as that of general education teachers: 32%
- Special educators are required to have knowledge, skills, and expertise that general education teachers are not: 84%

Measuring teacher effectiveness is impacted by:

- What is valued
- Our technological advances and limitations
- The data, evidence, and information we have or can acquire
- The resources (staff, money, time, policy levers) available to us and those we are willing to allocate to the task
Teacher effectiveness measured by student growth (e.g. value-added scores)

- Could devalue other ways teachers contribute to the growth and well-being of the student if not balanced with other measures
- Doesn’t identify effective practice and/or target professional growth in all circumstances
- A dearth of valid, reliable measures in many areas such as K-2, Art, etc.
Cutting-Edge Measurement—National Perspective

- Value-added analysis as a core component
- Softening measurement error through multiple measures
- Embedding measures in human capital decisions
- Correlating robust evaluation results with other measures
- We will discuss some promising practices later in the presentation
# Same vs. Differentiated System

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurements and dimensions identical</td>
<td>Lack of focus on social and behavioral outcomes</td>
</tr>
<tr>
<td></td>
<td>Difficult to measure progress for students on alternate standards</td>
</tr>
<tr>
<td></td>
<td>Fails to measure other roles/responsibilities (e.g. IEP facilitation, paperwork &amp; timelines, collaboration with families, supervising paraprofessionals etc.)</td>
</tr>
<tr>
<td>More simplistic</td>
<td>Could devalue the roles and responsibilities of special educators</td>
</tr>
<tr>
<td></td>
<td>Doesn’t encourage the use of evidence-based practices for students with disabilities</td>
</tr>
<tr>
<td>Fair</td>
<td>Not fair (e.g. poor general education instruction; limited role, fails to measure critical competencies; growth trajectory)</td>
</tr>
<tr>
<td>Increased Reliability</td>
<td>Evaluators may lack content/specialty knowledge</td>
</tr>
<tr>
<td></td>
<td>Lack of explicit criteria for specialty area teachers</td>
</tr>
</tbody>
</table>
Focus groups of special educators

- Determine if current system can assess special education teacher effectiveness or a need to differentiate the system
- Identify the challenges within the current system
- Identify instructional strategies, roles and responsibilities unique to special educators
- Identify standard measures to assess student growth
- Continually assess the effectiveness of the measure
Opinions Regarding Special Education Teacher Evaluation

- **92%**  
  *Special educator use of evidence-based strategies should be a component of the evaluation process.*

- **32%**  
  *Special educators should be evaluated using the same evaluation process as that of general education teachers.*

- **84%**  
  *Special educators are required to have knowledge, skills, and expertise that general education teachers are not.*
Battelle for Kids’ Philosophy

Balance is the key:

- Multiple data sources/measures
- Across time
- Linking teaching instruction to students
- Informing instruction based on data
- It’s about improvement, not judgment
What creative efforts are underway to measure effectiveness?

- Alternative ways to measure growth
  - Teacher Evaluation Advisory Committee (TN)
  - Austin Independent School District (TX): REACH
  - District of Columbia Public Schools: IMPACT
  - Harrison School District Two (CO)
  - Hillsborough County Public Schools (FL)
  - Gallup Student Poll (Various)
What are the basics of value-added?

- Value-added is a statistical measure designed to measure the growth of students vs. a growth standard.
- Value-added is typically based on standardized test scores.
- Value-added often uses prior test scores and makes a prediction about how a student will score in the future based on those scores.
- There are multiple providers of value-added data.
- Some models are very robust and some are relatively straightforward.
- Value-added is considered by many to be a “productivity” measure.
Yes, within reason

- No snap judgments, look across time
  - Repeatability reaches .8 with a three-year average
- Balance the value-added data with other measures
  - All statistical models are wrong but some are very useful
- Beware of unclean data… education data is notoriously noisy, we are now trying to use educational data systems for purposes for which they were not designed
- Be a smart consumer of data… sample size matters
- Simple models vs. complex models… you make tradeoffs and to a great degree it depends on the end use of the data
  - “Simpler is better unless it is wrong”
What does value-added attempt to measure?
Why do some see it as the “great equalizer”? 

The image shows a bar chart with the y-axis labeled as “Percent of Districts” and the x-axis labeled with categories: Rural Poor, Other Rural, Suburb, All Urban, and Urban 21. The bars are colored in green, red, and yellow, with different shades indicating the percentage distribution across these categories.
## Value-added at the district level

### 2007 Value Added Summary Report for School District 0
#### Mathematics

<table>
<thead>
<tr>
<th>School Name</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psi Elementary School</td>
<td>-0.6</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rho Middle School</td>
<td>-</td>
<td>-</td>
<td>-5.7</td>
<td>5.6</td>
<td>-1.6</td>
</tr>
<tr>
<td>Saturn Elementary School</td>
<td>1.8</td>
<td>1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigma Middle School</td>
<td>-</td>
<td>-</td>
<td>-10.5</td>
<td>2.9</td>
<td>-2.9</td>
</tr>
<tr>
<td>Tau Elementary School</td>
<td>1.9</td>
<td>-5.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theta Elementary School</td>
<td>-1.4</td>
<td>0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upsilon Elementary School</td>
<td>2.4</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xi Elementary School</td>
<td>4.9</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Progress significantly Above the average school in the testing pool.**
- **Progress Not Detectably Different from the average school in the testing pool.**
- **Progress significantly Below the average school in the testing pool.**

---

The school does not have data for this test and subject in the most recent year.
Value-added at the diagnostic level

![Graph showing gain by prior-achievement subgroups]

<table>
<thead>
<tr>
<th>Mathematics</th>
<th>Reference Line</th>
<th>Prior-Achievement Subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1 (Lowest)</td>
</tr>
<tr>
<td>Mathematics</td>
<td></td>
<td>0.0</td>
</tr>
<tr>
<td>2007</td>
<td>Gain</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>Std Err</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>No. of Students</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>% of Students</td>
<td>9.8</td>
</tr>
<tr>
<td>Previous Cohort(s)</td>
<td>Gain</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>Std Err</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>No. of Students</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>% of Students</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Value-added at the teacher level

SAS® EVAAS® Teacher Diagnostic Report for 2004

<table>
<thead>
<tr>
<th>2004</th>
<th>Observed minus Predicted Score by Predicted Score Subgroup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (Lowest)</td>
</tr>
<tr>
<td>Mean</td>
<td>−2.2</td>
</tr>
<tr>
<td>Std. Error</td>
<td>2.7</td>
</tr>
<tr>
<td>Nr of Students</td>
<td>9</td>
</tr>
<tr>
<td>% of Students</td>
<td>39.1</td>
</tr>
</tbody>
</table>
Why take the analysis to the teacher level?

- District Effects: 5%
- School Effects: 30%
- Teacher Effects: 65%
Challenges in Using Growth Models for Special Educators

- A research-derived value-added model for special educators does not exist
- Student learning trajectory
- Students assessed with accommodations
- Students assessed on alternate standards
- Small student samples commonly associated with special education caseloads
- Student mobility
- Teacher attribution in a coteaching situation
Opinions Regarding Use of Student Achievement for Special Educators

- **Achievement gains should be a component:**
  - Strongly Agree: 73%
  - Agree: 60%
  - Neither Agree nor Disagree: 21%

- **Standardized test scores should be a component:**
  - Strongly Agree or Agree: 60%
  - Neither Agree nor Disagree: 21%

- **Progress on the IEP should be a component:**
  - Strongly Agree: 73%
Opinions Regarding Attribution in Coteaching Setting

- **Both teachers held accountable for all students**
  - Strongly Agree or Agree: 85%
  - Neither Agree nor Disagree: 10%
  - Disagree: 5%
  - Strongly Disagree: 0%

- **Special educators in a coteaching role held accountable for all students**
  - Strongly Agree or Agree: 75%
  - Neither Agree nor Disagree: 15%
  - Disagree: 5%
  - Strongly Disagree: 0%

- **Special educators in a coteaching role held accountable for students with disabilities only**
  - Strongly Agree or Agree: 13%
  - Neither Agree nor Disagree: 45%
  - Disagree: 30%
  - Strongly Disagree: 12%

What criteria have been used?

- Determining “n” is really a local decision. We have worked with district that use small as “5” and as large as “10”.
- What trade offs are you willing to make?
  - Smaller “n” brings in more folks
  - Smaller “n” drives up standard error
  - Larger “n” excludes more folks
  - Larger “n” brings more “certainty”
The devil is in the details…

You cannot ask professionals to make significant changes in their practice unless they believe the data they are receiving is accurate and measures what it purports to measure.

- Mobility
- Co-teaching
- Data accuracy
- Data transparency
BFK•Link™ – Getting Started

Principal Set-Up Period

Buckley, Florence (10239)

- My Linkage
  - No Classes
  - Add Class

Manage Linkage: Cummings Middle (LD1_102)

- Setup
  - Review Admin Linkage How To Guide
  - Add/Remove Staff
  - Manage Linkage Support Team
  - Manage Linkage Setup
  - Review Student Linkage

Alerts and Notifications - Last run on 4/1/2010 1:13 PM

Linkage Completion Alerts
- Rosters with Zero Percent Linkage: 71 Alerts
- Rosters with no Students: 0 Alerts
- Grade/Subjects with low linkage: 13 Alerts
- Overclaimed Students: 0 Alerts

Review & Approve
- Review Linkage Completion Report
- Review Student/Subject Linkage Report
- Review Deleted Students
- Submit Final Linkage Approval

Key Linkage Periods

- Principal Set-Up
  - Ends in 3 days

- Teacher Linkage
  - 4/10/2010 - 4/23/2010
  - Begins in 3 days

- Principal Approval
  - 4/24/2010 - 5/7/2010
  - Begins in 17 days

Support Team
- Buckley, Florence (principal)

Help
- Staff Linkage How To Guide
- Submit Support Ticket

Admin Only
- Admin Linkage How To Guide
- Admin Linkage Checklist
- Staff Linkage Presentation
- Submit Support Ticket
Manage Your Linkage Support Team

Support team members can manage (add/remove) staff, staff classes and rosters, and modify linkages as necessary.

<table>
<thead>
<tr>
<th>Staff Name</th>
<th>E-mail Address</th>
<th>Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austin, Sherry</td>
<td><a href="mailto:Austin.Sherry@DemoDistrict.org">Austin.Sherry@DemoDistrict.org</a></td>
<td></td>
</tr>
<tr>
<td>Barber, Kurt</td>
<td><a href="mailto:Barber.Kurt@DemoDistrict.org">Barber.Kurt@DemoDistrict.org</a></td>
<td></td>
</tr>
<tr>
<td>Bentley, Christian</td>
<td><a href="mailto:Bentley.Christian@DemoDistrict.org">Bentley.Christian@DemoDistrict.org</a></td>
<td></td>
</tr>
<tr>
<td>Buckley, Florence (principal)</td>
<td><a href="mailto:Buckley.Florence@DemoDistrict.org">Buckley.Florence@DemoDistrict.org</a></td>
<td>✔️</td>
</tr>
</tbody>
</table>
Teacher Linkage

Teachers begin the linkage process by viewing a list of all their classes for tested subjects requiring linkage.

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Students</th>
<th>Building</th>
<th>Status</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integrated English</td>
<td>13</td>
<td>Bradshaw Elem</td>
<td>Not Started</td>
<td>Begin Copy Roster</td>
</tr>
<tr>
<td>Language Arts 4-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mathematics 4-6</td>
<td>21</td>
<td>Bradshaw Elem</td>
<td>Not Started</td>
<td>Begin Copy Roster</td>
</tr>
<tr>
<td>Science 4-6</td>
<td>18</td>
<td>Bradshaw Elem</td>
<td>Not Started</td>
<td>Begin Copy Roster</td>
</tr>
<tr>
<td>Social Studies 4-6</td>
<td>18</td>
<td>Bradshaw Elem</td>
<td>Not Started</td>
<td>Begin Copy Roster</td>
</tr>
</tbody>
</table>

Key Linkage Periods

- **Principal Set-Up**
  - Closed

- **Teacher Linkage**
  - 4/1/2010 - 4/23/2010
  - Ends in 17 days

- **Principal Approval**
  - 4/24/2010 - 5/7/2010
  - Begins in 17 days

Support Team

- Green, Edwin (principal)
Class rosters must be reviewed and accurately completed.
Percentage of Instruction (Mobility)

Mobility information is combined with % of instruction information. In Advance Mode, nine separate months of instruction are collected.
Setting Percentage of Instruction

- In most cases, teachers will set all students at 100%.
- Team teaching situations might share student instruction at 50% each.
- A special education teacher may claim a student for as little as 20%.
Student Learning Objectives:

- Teachers determine two SLOs for the semester/year
- One SLO must address all students, other may be targeted
- Use broad array of assessments
- Assess student needs more directly
- Align classroom, campus, and district expectations
- Aligned to state standards/campus improvement plans
- Based on multiple sources of student data
- Assessed with pre and post assessment
- Targets of student growth
- Peer collaboration

Slide courtesy of Laura Goe, ETS
# Rubric for student learning objectives

<table>
<thead>
<tr>
<th>LEVEL 4</th>
<th>LEVEL 3</th>
<th>LEVEL 2</th>
<th>LEVEL 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPROVAL</td>
<td>APPROVAL</td>
<td>APPROVAL</td>
<td>APPROVAL</td>
</tr>
</tbody>
</table>

Indicates expert use of data, rigorous goal-setting based on student and community strengths and needs, incorporates appropriate valid/reliable assessments, considers teacher’s own prior performance, and demonstrates alignment with campus improvement goals.

Indicates good use of data and acceptable level of rigor for students, considers teacher’s prior performance, and utilizes appropriate assessments, but is not aligned with broader campus goals and does not incorporate the strengths of students or the school community.

Suggests superficial use of data, is minimally rigorous for students and the teacher, uses related but unproven assessments, and does not incorporate the goals of the campus or the strengths of the school community.

Represents ineffective use of data, is not rigorous for students or the teacher, does not support the campus goals, utilizes inappropriate assessment, and does not consider strengths of students or the school community.

<table>
<thead>
<tr>
<th>NEEDS REVISION</th>
<th>NEEDS REVISION</th>
<th>NEEDS REVISION</th>
<th>NEEDS REVISION</th>
</tr>
</thead>
</table>

## Needs Assessment/Rationale:

**What are the needs?**

- Campus data are reviewed for areas of strength and need (within subject area, within grade level, within student group, examining the Campus Improvement Plan (CIP), etc.).
- Classroom data are reviewed for areas of strength and need (by subject area, by student group, by concepts/skills/behaviors).
- Classroom data are reviewed for areas of need, but needs of specific student groups are not examined in depth, and strengths are not identified.
- Classroom data are not used to identify student needs.
- Campus needs and strengths are not incorporated.

## Learning Content/Context and Objectives:

**Targets the needs of the identified population**

- Targets specific academic concepts, skills, or behaviors based on TEKS/TAKS Objective
- Targets the needs of the identified population.
- Targets year-long (or semester-long) concepts, skills, or behaviors.
- CIP Goals are not incorporated.

- Targets specific academic concepts, skills, or behaviors based on TEKS/TAKS Objective
- Targets the needs of the identified population.
- Targets year-long (or semester-long) concepts, skills, or behaviors.
- CIP Goals are not incorporated.

- Does not target concepts, skills, or behaviors based on TEKS/TAKS Objective
- Does not target year-long (or semester-long) concepts, skills, or behaviors.
- Does not target the needs of all students in the identified population.
- Does not consider the strengths of the identified population, classroom, or school community.
- CIP Goals are not incorporated.
<table>
<thead>
<tr>
<th>Rubric for student learning objectives (cont’d)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Objective</strong>: What will students learn?</td>
</tr>
<tr>
<td><strong>Level 4</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Based on the identified student needs.</td>
</tr>
<tr>
<td>Supports goals of the CIP.</td>
</tr>
<tr>
<td>Is a good example of ongoing, reflexive practice.</td>
</tr>
<tr>
<td>Provides clear focus for instruction and assessment.</td>
</tr>
<tr>
<td>Is measurable.</td>
</tr>
<tr>
<td>Reflects strengths of students and school community.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Outcome Assessment</strong>: How will you know whether they learned it?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 4</strong></td>
</tr>
<tr>
<td>Aligns with the targeted learning content area.</td>
</tr>
<tr>
<td>Relationship with learning objective is apparent.</td>
</tr>
<tr>
<td>Has been demonstrated as reliable and valid for targeted students.</td>
</tr>
<tr>
<td>Follows guidelines for appropriate assessments.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Performance Target</strong>: What is your goal for student achievement?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 4</strong></td>
</tr>
<tr>
<td>Is a rigorous expectation for students.</td>
</tr>
<tr>
<td>Is a rigorous expectation for teachers, based on past performance.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Rigor</strong>: How rigorous is your SLO?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 4</strong></td>
</tr>
<tr>
<td>Content is challenging, complex and progressively deepens knowledge of core content.</td>
</tr>
<tr>
<td>Content is thought-provoking requiring high thinking demand</td>
</tr>
<tr>
<td>Requires analytical thinking and active use of knowledge</td>
</tr>
<tr>
<td>Content is relevant to life/ experiences.</td>
</tr>
</tbody>
</table>
SLO Model Strengths/Weaknesses

Strengths

- Teachers take an active role in determining student learning goals
- Good professional growth opportunity for teachers
- If objectives are of high-quality and teachers plan instruction to meet them, students should benefit

Weaknesses

- Heavily dependent on administrator understanding and time commitment to supervision
- Not “comparable across classrooms” because teachers set the objectives and they will vary widely
- Not clear how students’ beginning point is determined

Slide courtesy of Laura Goe, ETS
Norwell Public Schools, Massachusetts

- Progress on the IEP is factored into evaluation of special educators.
- Each student has a data binder
- Teachers and administrators were trained extensively

Both SLO and IEP use to determine student growth are heavily dependent on teacher training and support.
Delaware Model

- Standardized test will be used as part of teachers’ scores in some grades/subjects
- “Group alike” teachers, meeting with facilitators, determine which assessments, rubrics, processes can be used in their subjects/grades (multiple measures)
- Assessments must focus on standards, be given in a “standardized” way, i.e., giving pre-test on same day, for same length of time, with same preparation
- Teachers recommend assessments to the state for approval
- Teachers/groups of teachers take primary responsibility for determining student growth

Slide courtesy of Laura Goe, ETS
Other Evidence of Teacher Effectiveness

- **Evidence of instructional quality**
  - Classroom observations (e.g. Charlotte Danielson’s Teaching and Learning Framework)
  - Lesson plans, assignments, and student work

- **Other evidence**
  - Administrator/supervisor reports
  - Surveys of students and/or parents
  - An “evidence binder” created & presented by the teacher
Observation Protocols

- **51%** Align to the state’s professional teaching standards.
- **85%** Use the same observation instrument as that of general education teachers.
- **12%** Use a modified or different observation instrument.
- **26%** Didn’t know.

“*Our evaluation tool was developed in the district over 40 years ago.*”

“*Our current evaluation system is outdated and applied to nothing.*”
Observation Protocol
Practical Example

➢ Alabama Department of Education’s Professional Education Personnel Evaluation Program

• Slightly modified for the following:
  ▪ Specialty area systems (speech paths, library specialist)
  ▪ Teachers of students with significant cognitive disabilities

• Competencies added in certain areas (e.g., classroom is expanded to include community settings, and academic content is expanded to include functional life skills.)
Expert Opinions Regarding Evaluators

- Require specialized training: 77%
- Require training for evaluators: 61%
- Should have experience in special education: 60%
- Require specialized training: 12%

Strongly Agree or Agree

- Require specialized training: 77%
- Require training for evaluators: 61%
- Should have experience in special education: 60%
- Require specialized training: 12%

www.BattelleforKids.org
Practical Examples

➢ Toledo’s Peer Assistance and Review
  • School-based teams evaluate.

➢ Norwell Public Schools, Massachusetts
  • All teachers are evaluated using the same instrument.
  • Two formative assessments are conducted:
    ▪ One with principal
    ▪ One with special education administrator
  • Each evaluator focuses on expertise areas.
  • Both work collaboratively to develop summative evaluation.
District of Columbia IMPACT
Teacher Categories

Teachers sorted according to teacher type:

- Group 1: Teachers for whom value-added data can be generated
- Group 2: Teachers for whom value-added data cannot be generated
- Group 3: Special education teachers
- Group 4: Non-itinerant English Language Learner (ELL) teachers and bilingual teachers
- Etc...20 different groupings
Practical Example: District of Columbia IMPACT

Special Education

- Individual Teacher Value-Added Scores
- Non-Value-Added Achievement 10%
- Teaching and Learning Framework 50%
- Commitment to the School 5%
- School Value-Added Scores 5%
- Core Professionalism
- IEP Quality Plan 15%
- IEP Timeliness 15%
Policy and Practice Recommendations

➢ Include special education administrators when revamping/designing evaluation frameworks.

➢ Identify a common framework that defines effective teaching for all teachers, differentiating for special educators as appropriate.

➢ Integrate evidence-based practices for students with disabilities into evaluation models.
Policy and Practice
Recommendations

- Improve data quality.

- In addition to—or, in some situations, in the absence of—appropriate standardized assessment data, incorporate other reliable evidence of teachers’ contributions to student learning into the teacher evaluation system, such as progress toward accomplishing IEP objectives and student learning objectives across broad academic and behavioral domains.
Policy and Practice Recommendations

- Ensure that evaluator training includes explicit training for evaluators of special educators and/or consider establishing a model of peer-to-peer observations or a model in which evaluators are matched to specific disciplines.

- Collaborate with teacher preparation programs to ensure that evidence-based practices are incorporated into teacher preparation coursework and professional development activities.

References


References


Lynn Holdheide, Vanderbilt University
P: 615-322-8150
E-Mail: lynn.holdheide@vanderbilt.edu

Tony Bagshaw, Battelle for Kids
P: 513-309-5652
E-Mail: tbagshaw@battelleforkids.org