

# Forum on Evaluating Educator Effectiveness: Critical Considerations for Including Students with Disabilities

2012



**Forum on Evaluating Educator Effectiveness:  
Critical Considerations for Including Students with Disabilities**

A joint publication of:

NATIONAL CENTER ON EDUCATIONAL OUTCOMES

The Center is supported through a Cooperative Agreement (#H326G110002) with the Research to Practice Division, Office of Special Education Programs, U.S. Department of Education. Additional support for targeted projects, including those on English language learners, is provided by other federal and state agencies. The Center is affiliated with the Institute on Community Integration at the College of Education and Human Development, University of Minnesota. Opinions expressed herein do not necessarily reflect those of the U.S. Department of Education or Offices within it.

Assessing Special Education Students (ASES)  
State Collaborative on Assessment and Student Standards (SCASS)  
THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

All rights reserved. Any or all portions of this document may be reproduced and distributed without prior permission, provided the source is cited as:

Warren, S., Thurlow, M., Lazarus, S., Christensen, L. Chartrand, A., and Rieke, R. (2012). *Forum on evaluating educator effectiveness: Critical considerations for including students with disabilities*. Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes and Washington, DC.: Council of Chief State School Officers, *Assessing Special Education Students State Collaborative on Assessment and Student Standards*.

Available on the World Wide Web at <http://www.nceo.info> and [http://www.ccsso.org/Resources/Programs/Assessing\\_Special\\_Education\\_Students\\_\(ASES\).html](http://www.ccsso.org/Resources/Programs/Assessing_Special_Education_Students_(ASES).html).

### **Forum Participants**

Heather Buzick (Educational Testing Service)  
Wendy Carver (Utah State Office of Education)  
Anne Chartrand (Southeast Regional Resource Center)  
Beth Cipoletti (West Virginia Department of Education)  
Jason Clyuner (Colorado Department of Education)  
Andrew Hinkle (Ohio Department of Education)  
Lynn Holdheide (Vanderbilt University)  
Greg Keith (Minnesota Department of Education)  
Sarah Kennedy (Keystone Assessment Group)  
Judy Kraft (University of Kentucky)  
Tammy Lawson (Oklahoma Department of Education)  
Deborah Matthews (Kansas State Department of Education)  
Daniel Mix (Curriculum Associates)  
Caroline Mock (NWEA)  
Terrence Moore (Ohio Department of Education)  
Debra Price-Ellingstad (Minnesota Department of Education)  
Pat Reynolds (Arizona Department of Education)  
Christie Stephenson (Oklahoma Department of Education)  
Wendy Stoica (Ohio Department of Education)  
Catherine Taylor (University of Washington)  
Brian M. Touchette (Delaware Department of Education)  
Vicki Weinberg (Minnesota Department of Education)  
Daniel Wiener (Massachusetts Department of Education)  
Leila Williams (Arizona Department of Education)  
Debbie Willis (American Printing House for the Blind)  
Anonymous (Pacific Metrics)

### **Forum Facilitators**

Anne Chartrand (SERRC)  
Laurene Christensen (NCEO)  
Sheryl Lazarus (NCEO)

### **Forum Recorders**

Kamarrie Davis (NCEO)  
Jane Krentz (NCEO)  
Amanda Morse (NCRRC)

### **Forum Notetakers**

Lynn Price (NCEO)  
Rebekah Rieke (NCEO)  
Christopher Rogers (NCEO)

### **Forum Conveners**

Martha Thurlow (NCEO)  
Sandra Warren (ASES SCASS CCSSO)

## **Forum on Evaluating Educator Effectiveness: Critical Considerations for Including Students with Disabilities**

### **Background**

Thirty-six individuals representing 12 states, 4 testing companies, and 6 other organizations participated in a forum on June 26, 2012 in Minneapolis, Minnesota, to discuss educator effectiveness evaluation. The forum was a pre-session to the Council of Chief State School Officers (CCSSO) National Conference on Student Assessment, and was a collaboration of the Assessing Special Education Students (ASES) State Collaborative on Assessment and State Standards (SCASS) and the National Center on Educational Outcomes (NCEO).

### **Purpose**

The purpose of the forum on evaluating educator effectiveness was to address the challenging issues that exist in including students with disabilities and their teachers in educator evaluation systems. The primary goal of the forum was to provide state representatives and other interested parties with an opportunity to meet and share their perspectives on this topic, particularly in light of the changing nature of accountability systems across the United States. This topic is particularly important as the nation focuses on the development of assessment systems by the Race-to-the-Top Assessment Consortia, the Partnership for Assessment of Readiness for College and Careers (PARCC) Consortium and the SMARTER Balanced Assessment Consortium (SBAC), and as states apply for and receive flexibility for their accountability systems.

The specific focus of the forum discussion was three issues related to educator effectiveness evaluation systems:

1. Including assessment results of students with disabilities
2. Using IEP goals or student learning objectives (SLO)
3. Developing a balanced model for educator effectiveness evaluations

Each of these issues was defined for the participants and a set of guiding questions provided. The time for discussion was short and discussions were intense. The agenda for the four-hour forum was as follows:

- Welcome (Sandra Warren, CCSSO, and Martha Thurlow, NCEO)
- National Overview (Lynn Holdheide, National Comprehensive Center for Teacher Quality)
- Lessons Learned from States
  - Arizona (Leila Williams)
  - Delaware (Brian Touchette)
  - Minnesota (Greg Keith)
- Break-out Sessions (with short introductions)
  - How Should Assessment Results of Students with Disabilities be Used in Evaluation Models of Educator Effectiveness? (facilitated by Anne Chartrand, SERRC)
  - What are the Benefits and Concerns of Using IEP Goals or Student Learning Objectives (SLO)? (facilitated by Sheryl Lazarus, NCEO)
  - If Multiple Measures are Used, What Would a Balanced Model Look Like? (facilitated by Laurene Christensen, NCEO)
- Final Discussion and Plans for Next Steps

### **Structure of This Report**

Although this report summarizes the introductory information provided to forum participants, its main purpose is to describe the facilitated forum topic discussions themselves. Summaries for the discussions were developed from notes taken by recorders (providing comments on flip charts) and note takers (entering comments into computers).

This report first summarizes the introductory material provided to participants, and then presents a summary of the discussion in each topic session. Discussions in each topic session addressed the following two questions:

1. What are the issues and implications for the topic?
2. What do we need to do – in terms of research, policies, and resources?

Participants were encouraged to comment and discuss freely, with assurances that no individual's name, nor any state, company, or organization names would be attached to comments that were made. Complete anonymity of statements was assured. This led to frank and open conversations.

### **Session Introduction**

Dr. Sandra Warren, ASES SCASS Advisor, provided an overview of the forum and discussed the goal of discussing measures of student achievement related to evaluating educator effectiveness. This is an issue of increasing importance as virtually every state is revamping their system to meet required aspects of Race-to-the-Top funding and Principle 3 of the ESEA Flexibility Initiatives.

The U.S. Department of Education has addressed the need to include students with disabilities and their teachers in the educator effectiveness plans for their flexibility applications. During the first rounds of Peer Reviews, the Department noted “Peers had several general concerns regarding SEAs’ proposals to develop and implement teacher and principal evaluation and support systems...” They go on to state “...these include, for example, lack of specificity regarding the process for determining the validity and reliability of the evaluation measures and how they will be consistently applied across LEAs. Further peers had concerns related to the measures SEAs intended to use to evaluate teachers of non-tested grades and subjects.” In nearly all applications, special education teachers were considered to be part of the group, “teachers of non-tested grades and subjects.” Additional information on this topic may be found at <http://www.ed.gov/sites/default/files/considerations-strengthen.pdf>.

## **Setting the Stage for Topic Discussions**

Dr. Martha Thurlow, NCEO Director, set the stage for the topic discussions. She indicated that part of the goal of discussions was to raise questions, particularly ones that will help in thinking about the future of educator evaluation and the inclusion of students with disabilities and their teachers. She indicated that in the first session, participants would hear from Lynn Holdheide from the National Comprehensive Center for Teacher Quality on the results of a national view of educator evaluation systems, as well as the findings from a September 2011 forum on the topic.

Following Lynne's presentation, three states shared the status of their proposed education effectiveness evaluation systems. Each addressed where students with disabilities and their teachers fit in the system being described.

In this report, we provide brief summaries of each of the presentations. These are followed by summaries of the discussions of the three topics.

## National Overview

Lynn Holdheide with the National Comprehensive Center for Teacher Quality at the American Institutes for Research gave an overview of national efforts in evaluating teachers of students with disabilities. Ms. Holdheide noted that to date, not much emphasis has been placed on considering the unique roles and responsibilities of special educators as well as the challenge of measuring and using growth of students with disabilities in teacher evaluation design.

Lynn provided several dimensions for state consideration when designing teacher evaluation models:

- First, designers must address the unique challenges in accurately measuring growth of students with disabilities and connecting that growth to teacher effects. Little is known about using student growth as a component of teacher evaluation. This is the case for all students but even more so for students with disabilities—most particularly for students participating in the alternate assessment. Designers should consider how students with disabilities, and their teachers, fit into existing and/or potential measures of growth in teacher evaluation and think through considerations of implementation for students with disabilities (see *Using Student Growth to Evaluate Educators of Students with Disabilities: Issues, Challenges, and Next Steps: A Forum of State Special Education and Teacher Effectiveness and Researchers and the TQ Center Research & Policy Brief Measuring Teachers' Contributions to Student Learning Growth for Nontested Grades and Subjects*).
- Second, designers must consider how the various measures of instructional practice (e.g., observation protocols, student and parent surveys, and evaluation of artifacts) are appropriate for use with teachers of students with disabilities—or whether the field would benefit from the strengthening of the existing protocols that speak to specific evidence-based instructional practices for students with disabilities (e.g., direct and explicit instruction and learning strategy instruction), specific roles and responsibilities of special educators (e.g., individualized education program [IEP] facilitation, development, and implementation and coordination of related services personnel), and specific curricular needs (e.g., secondary transition services, social and behavioral needs, and orientation and mobility).
- Third, designers must contemplate the distinct considerations for teachers (both general and special education) serving in a co-teaching capacity. The chief consideration is how student growth will be accurately and fairly attributed when more than one teacher is contributing to student learning. Likewise, should measures of instructional practice be



modified, and are there indicators of effective co-teaching that should be factored when determining teacher effectiveness?

Lynn spent a significant portion of the presentation reviewing the challenges states face when using the most widely used approaches to measure student growth for the purpose of teacher evaluation (e.g. value-added, student learning objectives, curriculum-based measures, group-wide value added) and provided initial state considerations and needed areas of research to guide development and implementation efforts. Through interactive discussion, Lynn indicated that much work remains in this area and further suggested a national, strategic approach in conducting research and validating tools and implementation efforts so that more is known about effective frameworks for the evaluation of teachers of students with disabilities.

Making certain that the needs of students with disabilities and their teachers are fully represented within the design process from the very beginning is central to ensuring that the evaluation process leads to quality feedback regarding teacher performance. This feedback can be used to guide professional development and support leading toward improved teacher capacity and student learning—which is, it is to be hoped, the ultimate goal of evaluation systems.

## Lessons Learned from States

### Arizona (Leila Williams)

Leila Williams with the Arizona Department of Education (ADE) provided an overview of the Arizona Framework for Effective Teachers and Leaders. Arizona has already included in Statute:

*“on or before December 15, 2011 adopt and maintain a model framework for a teacher and principal evaluation instrument that includes quantitative data on student academic progress that accounts for between thirty-three percent and fifty per cent of the evaluation outcomes and best practices for professional development and evaluator training. School districts and charter schools shall use an instrument that meets the data requirements established by the State Board of Education to annually evaluate individual teachers and principals beginning in school year 2012 –2013.”*

ADE staff are working closely with local education agencies (LEAs) to provide guidance, strategies, and resources to:

- Improve the effectiveness of teachers and leaders
- Achieve equitable distribution of effective teachers
- Support targeted professional development based on identified needs
- Assist with the implementation of the Arizona Framework for Measuring Educator Effectiveness

Some of the challenges that districts are facing:

- Developing a framework for the group B teachers (non-assessed areas on state assessment)
- Identifying assessments in addition to state assessments that could be included
- Small LEAs/Charters with limited resources.
- Districts have mixed opinions about Special Education teachers (some are the Content teacher while others are not)

Leila shared how LEAs can weight groups (Group A – Content Teachers – Math/ELA and Group B non-assessed content):

Group A	Group B
Sample 1: <ul style="list-style-type: none"> <li>o 33% Classroom-level data</li> <li>o 17% School-level data</li> <li>o 50% Teaching performance</li> </ul> Sample 2: <ul style="list-style-type: none"> <li>o 50% Classroom-level data</li> <li>o 50% Teaching performance</li> </ul> Sample 3: <ul style="list-style-type: none"> <li>o 33% Classroom-level data</li> <li>o 67% Teaching performance</li> </ul>	Sample 1: <ul style="list-style-type: none"> <li>o 17% Classroom-level data</li> <li>o 33% School-level data</li> <li>o 50% Teaching performance</li> </ul> Sample 2: <ul style="list-style-type: none"> <li>o 50% School-level data</li> <li>o 50% Teaching performance</li> </ul> Sample 3: <ul style="list-style-type: none"> <li>o 33% School-level data</li> <li>o 67% Teaching performance</li> </ul>

### **Delaware (Brian Touchette)**

Brian Touchette with the Delaware Department of Education gave an overview of the status of the educator evaluation system in Delaware, as of the middle of June, 2012. He provided the website, which is where updated information on the system is provided ([www.doe.k12.de.us/csa/dpasii/default.shtml](http://www.doe.k12.de.us/csa/dpasii/default.shtml)). He noted that there are five parts to Delaware's educator evaluation system. It includes four components from Charlotte Danielson's work (planning and preparation, classroom environment, instruction, and professional responsibilities), plus growth measures of student improvement as a fifth component.

Brian identified the challenges of determining the educator groups to be evaluated and determining what measures to use to show student improvement, as well as how to match educators to measures. He noted that three educator groups had been identified: (I) reading or math teachers of record for at least 10 students who participate in the Delaware Comprehensive Assessment System (DCAS) in grades 3-10; (II) teachers of record who generally give grades for at least 10 students at any grade or content area other than grades 3 – 10 for reading or math; and (III) any educator who is not in Group I or II.

He also noted that three types of measures had been identified to examine student improvement. They include: A – state assessment data, B – external and internal measures, and C – growth measures. Each of these was explained in more detail. Readers are encouraged to go to the Delaware website for more details. Brian also noted that for the growth targets for the state assessment, students will be compared to “similar students,” which include: those in the same

grade and content area, those with the same fall score, and those who were the same in terms of their status as English language learners (ELL), students with disabilities (SWD), English language learners with disabilities (ELLwD), and non-ELL/SWD.

Finally, Brian indicated that the three types of educator groups would be evaluated against the different types of measures in different proportions. For example, educators in group I would be 50% on the state assessment and 50% on the external/internal measures. Educators in group II, would be 50% external/internal and 50% growth measures. Group III educators would be measured 100% with growth measures.

### **Minnesota (Greg Keith)**

Greg Keith of the Minnesota Department of Education shared the work of Minnesota in developing an educator evaluation system. Greg noted that Minnesota is still early in the process of developing the evaluation system, which was required by legislation approved in 2011. Minnesota is a local control state, and as a result, districts are permitted to develop their own model or use the state model. In either case, the evaluation system must be based on the Minnesota teaching standards and be tied to professional development, including professional learning communities.

Greg identified the key components of the system to include multiple measures over a three-year cycle, including observations of teaching, one summative assessment, and a portfolio that may include samples of video. The evaluation system uses value-added for areas in which the students are assessed for accountability, and uses student growth as a measure in the non-assessed areas.

With regard to students with disabilities and English language learners, Greg mentioned a number of strategies that Minnesota has implemented. Educators in both the areas of special education and English language acquisition have been convened to consider the inclusion of special populations. This group has recommended that inclusion be part of the system as a non-tested component that also includes both group-wide SLOs and targeted SLOs based on student need. A collaborative process is needed in order to identify assessment performance levels.

Greg identified the challenges to include the impact of the political process and the generally short timeline. In Minnesota, there had been a push to base teacher layoffs and tenure on the evaluation system; the state has approached this challenge by getting stakeholders (i.e., teachers and other educators) to have ownership in the process. Minnesota has coped with developing this evaluation system within a short timeline by borrowing ideas from other states, seeking outside expertise, and engaging districts in the process. Minnesota hopes, Greg notes, to maintain a long-term vision for educator effectiveness.

Greg emphasized the importance of careful implementation. He mentioned that Minnesota has ensured a strong design for implementation, and that evaluators must be well-trained for the system to be most effective. The system will be piloted in Spring of 2013.

## **How Should Assessment Results of Students with Disabilities be Used in Evaluation Models of Educator Effectiveness?**

This discussion centered on ways to develop teacher evaluation systems that will appropriately include assessment results of students with disabilities. Participants recognized that there are some complex issues that need to be considered—but stressed that it is important to include all students right from the start rather than trying to retrofit teacher evaluation systems.

### **Question 1: What are the issues and implications regarding how assessment results of students with disabilities should be used in evaluation models of teacher effectiveness?**

As states and districts move to the implementation of teacher evaluation systems, there are many wonderful opportunities (as well as some challenges) related to the inclusion of the results of students with disabilities in teacher evaluation systems.

#### ***Theme 1 – Teacher evaluation systems should include special education teachers and their students.***

- A benefit of new requirements (e.g., ESEA flexibility waivers, Race-to-the-Top funding) to develop teacher evaluation systems is that they will provide an opportunity for the special education voice to insert itself into the conversation and be heard. States are not experts but they are putting out what they think is their best thinking and they are bringing everybody to the table to develop measures that may be imperfect together. Special education needs to seize this opportunity to work with general education and “get on the radar” even more than before.
- Teacher evaluation systems will provide an opportunity to really see “what is happening” so it is vital to include students who receive special education services and as well as other subpopulations. There is a high expectation and belief that in paying attention to students with disabilities we will see movement and growth. The lowest performing students are not just kids with disabilities, but they are struggling and we are noticing them.
- Both general and special education teachers have a role in educating students with disabilities, and general education teachers need to care about accessibility and ensure that all students can show what they now.
- Participants touched on the benefit of professional development for teachers and administrators that will result from new requirements to evaluate teacher effectiveness that include results for all students. One participant shared a result of his state’s work to

develop a teacher evaluation system that includes results of all students: “The PD that teachers get is incredible because they are actually learning how to build a test.”

***Theme 2 - If growth models are used in a state, students who receive special education services should be included in the same way as other students.***

- Participants stressed that growth needs to be measured well for all students (including students with disabilities), and warned against possible negative consequences that poor measurements could have on both students and teachers. For interim types of measures, how do we keep people from lauding growth that is not actually being measured correctly? The pretest needs to be a valid measure of student performance. Growth needs to be validated by other measures to make a clear picture. Most of these issues are not specific to special education.
- Care needs to be taken so that growth models do not end up having the unintended consequence of putting students in more restricted environments.
- Growth models should be selected that support inclusion of the alternate assessment. For example, one state’s alternate assessment does not include scale scores but the growth model includes them. Alternate assessments without scale scores are still a real challenge in a value-added model.

***Theme 3 - Teacher evaluation systems should be flexible and inclusive of alternate assessments results.***

- A flexible system should be built from the beginning: “We are building this as we go” and it should include students who participate in alternate assessments. States do not have the luxury to wait for research to happen. So, the challenge is to design a system that is constantly being evaluated and open to change as it is developed. States need to know what questions they want to have answered as they go about designing it and be clear about those questions.
- Cohort sizes need to be considered when thinking about models that will work for teachers who have small numbers of students. One participant communicated that her state has struggled with cohort sizes when considering a teacher evaluation system, and as a result special education students are often “dropped” from the results and not included. The average class size in this state is between 6-10 students for alternate assessments and such small numbers make it difficult to determine what is going on because there are rarely over 15 special education students still taking the general assessment. As a result, special education was excluded: “Twenty five students is the ideal, we’ll allow it to go down to 10 with administrative approval. If it is under 10

students, then we have to beef it up with more data.” For some kids “maintaining” would be considered progress.

***Theme 4 - As the Common Core State Standards (CCSS) are being rolled out there is uncertainty about how students with disabilities will have access to academic content, be able to show what they know on assessments, and consequently, whether their results will contribute to a fair evaluation of their teachers.***

- There are still many unknowns about how the curriculum will look for students with disabilities as a result of the CCSS. All students need access to content.
- The implementation of CCSS should not result in students with disabilities being excluded from teacher evaluation systems due to a misguided perception that only “typical” students should be included in teacher evaluation systems.
- One participant reminded others of the ultimate goal to keep in mind as the CCSSs are rolled out: “If we can agree it’s about the kids and come back to this over and over again,” then we will stay on track. “The formula is complicated and people are so panicked that they don’t understand or remember what the ultimate goal is.”

**Question 2: What do we need to do – in terms of research, policies, and resources?**

There is a need for more research to help us develop a better understanding of how to appropriately include the results of students with disabilities in teacher evaluation systems. There is also a need to develop communications plans so that stakeholders have a better understanding of how special education teachers and their students are included in the teacher evaluation systems.

***Theme 1 – There is an underlying need to explain to the public, parents, and teachers what new models for teacher evaluation mean for all teachers, including special education teachers and their students.***

- There is a need to develop a clear way for each state to explain what is happening around the evaluation of special education teachers, with clear explanations that are non-technical. States also need to communicate how students who receive special education services are included in the evaluation of general education teachers who teach students who receive special education services. A participant said, “We are terrible at PR and in the newspaper only when it is bad.”
- Communication with teachers is important because teachers need to understand why professional development related to teacher evaluation matters and “have got to buy into what is happening.” A fear was expressed that most teachers think that “value-added”



and growth are “all that matters” and that providing content access to students is forgotten.

***Theme 2 – Decisions need to be made about who exactly is evaluated and what exactly the evaluation system will look like.***

- In principal evaluation systems, students who receive special education services should be included in the model the same as other students so as to be aware of students with disabilities in their schools.
- Because many of the teacher evaluation systems are at the district level, it is important that states have appropriate frameworks that will help districts to do the right thing. A participant said, “Evaluation systems could vary not only from district to district but from room to room.”
- One participant noted that nobody is questioning “who is the teacher,” pointing out that every state has different types of teachers (e.g., long-term subs, “Teach for America” volunteers, etc.). The conversation could then move to “who gets out of being evaluated?” The data we are using ought to point us back to what the ingredients are that are giving us certain results. If we want to know more we want to be able to look at whatever feeds into data that says we are successful: “The evaluation of this constellation of things helped us get here.” So we have to be able to map back to all the contributors.

## **What are the Benefits and Concerns of Using IEP Goals or Student Learning Objectives (SLOs)?**

This discussion centered on considerations related to the use of IEP Goals or Student Learning Objectives (SLOs) in teacher evaluation systems. The initial discussion focused on the use of IEP Goals followed by a lengthier discussion of SLOs.

### **Question 1: What are the issues and implications regarding the benefits and concerns of using IEP Goals or Student Learning Objectives (SLOs)?**

Session participants did not believe that IEP goals were appropriate to use in teacher evaluation systems, so the discussion focused on SLOs. Session participants generally believed that there were many challenges to using SLOs for teacher evaluation purposes, and questioned whether those challenges can be resolved.

***Theme 1- An unintended consequence of using SLOs as a measure for teacher evaluation may be lowered expectations for students with disabilities.***

- Concern was expressed about forcing teachers to participate in this system and then “putting them on the spot” by making them responsible for each student’s success relative to the measures. More questioning needs to happen as to whether there is good justification for using SLOs as the core for evaluating teachers. Participants cautioned against the use of SLOs to measure teacher evaluation thereby making it a very “high-stakes” scenario for teachers.
- If SLOs are used they should only be a small component of the overall “score” that a teacher receives. One participant shared that in his state, SLOs are not the only measuring tool used to evaluate teachers. The entire evaluation system includes much more than SLOs, which comprise 10% of the evaluation piece. As a result, SLOs are utilized as one growth measurement among many, complementing the other 80% of measurement.
- One participant expressed a concern about how the process of developing SLOs can have the effect of lowering the final objectives so that students can achieve a good score in order to then show improvement. Other participants agreed that often, when we add high stakes to an evaluation, this kind of “low-balling” tends to happen.
- One state shared that about 90% of teachers who have developed SLOs have met them but the state is unsure how valid the SLOs actually are: “It’s not on a bell curve.” This state also has a challenge of having to measure growth over a two year period. Measuring from the beginning of one year to the end is easier than doing so over two years because of summer regression.

***Theme 2 - If SLOs are used teachers will need a lot of training and professional development to develop and use SLOs.***

- Most teachers do not have the general skills to write their own objectives and create SLOs and then assess students on these SLOs.
- A participant stressed that an important goal of teacher evaluation should include teacher improvement over time. There is a concern that the correct support is not being provided to teachers.

**Question 2: What do we need to do – in terms of research, policies, and resources?**

There is a need for research to get a better understanding of the advantages and challenges of using SLOs to measure student growth. If SLOs are used in teacher evaluation systems there will be a need for substantial professional development so that teachers learn how to use them well.

***Theme 1 - There is a need for the field to better understand how to use SLOs to measure student growth.***

- The use of SLOs for teacher evaluation is moving very quickly and participants expressed a need to slow things down. Several cautioned against moving ahead with using SLOs for teacher evaluation before reevaluating what is already known about how to measure growth.
- States and districts sometimes go back to old practices (even if they are not very good) because it is what they know. Many participants felt that while it has been decided by many in special education that evaluating teachers using SLOs is not a best practice, they wondered simultaneously what kinds of measures would be appropriate and worried that if good measures are not established, “states go back to what they know” and will use SLOs.

***Theme 2 - If SLOs are used, more work is needed to validate their use to capture the growth of students with disabilities.***

- Many assumptions go into the development of SLOs. SLOs seem to be a default option that states use, but this does not negate the need for developing other methods for measuring growth of students with disabilities. SLOs are an option because states have to come up with something, but there are many challenges with them. States need to focus in on determining and developing the rigor of SLOs.
- An evidence-based data collection system needs to be in place in order to authenticate a teacher evaluation system using SLOs, which in turn, may remedy some of the concerns discussed above. By collecting data and testing certain claims by using validity

principles, practices, technology and instruments, what is and what is not working may be clarified.

- In one state, one way to help build stronger SLOs is by having principals and other administrators validate the SLOs that are developed by the teachers. However, this assumes that these administrators know what they are doing and would not lower expectations.
- Researchers need to study SLOs in a methodological way and then provide this information to states and also provide states with a solution when something does not work. States do not always have the ability to defend the validity of what they are doing, but they also do not have any other options at this point.

***Theme 3 - Open communication with, and inclusion of teachers in, the process of SLO development is needed.***

- Participants agreed that there is a need for open respectful conversations about SLOs. We need to build trust in the field surrounding this system so that teachers “buy-in” to the system.
- One participant observed an assumption in the teaching community that teachers know everything they need to know about the students they teach and the materials that are available to them. In many cases, teachers may be hesitant to reveal that they do not know how to go about doing something, for fear of negative feedback. Instead, a respectful dialogue needs to take place about areas in which teachers need training.
- While mostly concerned about the use of SLOs as tools for evaluating teachers, participants acknowledged that one possible benefit could be that this would potentially engage teachers and include all teachers in a teacher evaluation system that would better serve all students, including students with disabilities.

## **If Multiple Measures are Used, What Would a Balanced Model Look Like?**

This discussion focused on the various measures that might be used in a teacher evaluation system, and how they might be combined into a balanced system.

### **Question 1: What are the issues and implications regarding the use of multiple measures in educator effectiveness models?**

The discussion of multiple measures focused broadly on the inferences that can be drawn from the various data sources that may be available in states. Concern was raised about drawing conclusions about the relationships between different sets of data, and also about inconsistencies related to how students with disabilities are included in instruction and assessment. Too many measures may lead to lack of clarity.

#### ***Theme 1—Some measures do a better job of including students with disabilities than others.***

- When multiple measures are used, they should all be consistent with the statewide assessments.
- District-level assessments are often considered for use in evaluation systems; however, some district-level assessments have limitations on the use of accommodations for students with disabilities. It is important for a system of educator evaluation to account for differences in assessment practices, in order to most appropriately include students with disabilities.

#### ***Theme 2—Instruction should figure into a system of multiple measures.***

- The system itself should be used to improve instruction, and so the system and the measures included should be considered with this in mind.
- Inclusion of too many measures may lead to a “watered down” system of evaluation; it is important to determine what exactly should be measured. For example, a system that focuses on high-level reading skills will not measure growth in the same way as a system that is measuring the direct instruction of reading.
- The opportunity to learn is also an important factor to consider as a measure in an educator evaluation system. How can such a system account for students who are removed from instruction?

**Question 2: What do we need to do – in terms of research, policies, and resources?**

There was strong agreement among the group that multiple measures is the most appropriate way to evaluate educator effectiveness; however, there was no real consensus on what specific measures should be included.

***Theme 1—Additional research is needed on measures that might be used.***

- One state requested more research that can support the writing of measureable IEP goals. This is a concern that has existed in the field, and not enough research is currently available. If attainment of IEP goals is potentially one measure in an evaluation system, there needs to be greater understanding and capacity for writing measureable goals.
- An additional research concern was related to the use of rating scales, and how raters are trained to use those scales. Given the stakes posed by evaluation systems, there needs to be greater assurance related to rating scale invariance.

***Theme 2—Critical components or measures in an evaluation system need further consideration.***

- Student exemplars that demonstrate target skills was put forth as one necessary component of an evaluation system. These exemplars could be critiqued as a means of improving instruction.
- Observations of teachers was generally considered to be a critical component. Some states present suggested that teacher observations could be completed by administrators. In other states, teacher observations were suggested to be done by other teachers in a professional learning community.
- One state noted that having a vision, or target for the evaluation and observation is important if observations are included as a measure. Another participant suggested that focusing on a narrow range of practices (e.g., no more than 6) allows for observations that can improve instruction.

## **Discussion**

The forum on evaluating educator effectiveness generated many interesting comments and indications that challenging issues still exist. The comments summarized in this document provide important fodder for states and Assessment Consortia as they revise, improve, and develop their policies and practices surrounding educator effectiveness and evaluation.

Thoughtfully considering, and then addressing, the issues raised by forum participants will help to ensure Consortia and states expand policy discussions to include educators working the complete range of students with and without disabilities.