

120. LOCAL DISTRICT EVALUATION POLICY.

Each school district board of trustees will develop and adopt policies for teacher performance evaluation in which criteria and procedures for the evaluation of certificated personnel are research based and aligned to Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. The process of developing criteria and procedures for certificated personnel evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the certificated personnel for whom it is written. (3-29-10)

01. Standards. Each district evaluation model shall be aligned to state minimum standards that are based on Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. Those domains and components include: (3-29-10)

- a.** Domain 1 - Planning and Preparation: (3-29-10)
 - i. Demonstrating Knowledge of Content and Pedagogy; (3-29-10)
 - ii. Demonstrating Knowledge of Students; (3-29-10)
 - iii. Setting Instructional Goals; (3-29-10)
 - iv. Demonstrating Knowledge of Resources; (3-29-10)
 - v. Designing Coherent Instruction; and (3-29-10)
 - vi. Assessing Student Learning. (3-29-10)
- b.** Domain 2 - Learning Environment: (3-29-10)
 - i. Creating an Environment of Respect and Rapport; (3-29-10)
 - ii. Establishing a Culture for Learning; (3-29-10)
 - iii. Managing Classroom Procedures; (3-29-10)
 - iv. Managing Student Behavior; and (3-29-10)
 - v. Organizing Physical Space. (3-29-10)
- c.** Domain 3 - Instruction and Use of Assessment: (3-29-10)
 - i. Communicating Clearly and Accurately; (3-29-10)
 - ii. Using Questioning and Discussion Techniques; (3-29-10)
 - iii. Engaging Students in Learning; (3-29-10)
 - iv. Providing Feedback to Students; (3-29-10)
 - v. Demonstrating Flexibility and Responsiveness; and (3-29-10)
 - vi. Use Assessment to Inform Instruction and Improve Student Achievement. (3-29-10)
- d.** Domain 4 - Professional Responsibilities: (3-29-10)
 - i. Reflecting on Teaching; (3-29-10)
 - ii. Maintaining Accurate Records; (3-29-10)
 - iii. Communicating with Families; (3-29-10)
 - iv. Contributing to the School and District; (3-29-10)
 - v. Growing and Developing Professionally; and (3-29-10)
 - vi. Showing Professionalism. (3-29-10)

02. Participants. Each district evaluation policy will include provisions for evaluating all certificated employees identified in Section 33-1001, Idaho Code, Subsection 13, and each school nurse and librarian (Section 33-515, Idaho Code). Policies for evaluating certificated employees should identify the differences, if any, in the conduct of evaluations for nonrenewable contract personnel and renewable contract personnel. (4-1-97)

03. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: (4-1-97)

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions. (4-1-97)

b. Evaluation criteria -- statements of the general criteria upon which certificated personnel will be evaluated. (4-1-97)

c. Evaluator -- identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation. (4-1-97)

d. Sources of data -- description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation should be included as one (1) source of data. (4-1-97)

e. Procedure -- description of the procedure used in the conduct of certificated personnel evaluations. (4-1-97)

f. Communication of results -- the method by which certificated personnel are informed of the results of evaluation. (4-1-97)

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel. (4-1-97)

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations. (4-1-97)

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action. (4-1-97)

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's personnel evaluation system. (4-1-97)

k. Professional development and training -- a plan for ongoing training for evaluators/administrators and teachers on the districts evaluation standards, tool and process. (3-29-10)

l. Funding -- a plan for funding ongoing training and professional development for administrators in evaluation. (3-29-10)

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. (3-29-10)

n. A plan for how evaluations will be used to identify proficiency and define a process that identifies and assists teachers in need of improvement. (3-29-10)

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and

administrators in the development and ongoing review of their teacher evaluation plan. (3-29-10)

04. Evaluation Policy - Frequency of Evaluation. The evaluation policy should include a provision for evaluating all certificated personnel on a fair and consistent basis. At a minimum, the policy must provide standards for evaluating the following personnel: (4-1-97)

a. First-, second-, and third-year nonrenewable contract personnel will be evaluated at least once prior to the beginning of the second semester of the school year. (4-1-97)

b. All renewable contract personnel will be evaluated at least once annually. (4-1-97)

05. Evaluation Policy - Personnel Records. Permanent records of each certificated personnel evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code). (4-1-97)

Initiative and Selection Criteria		Purpose
<p>The Idaho Mentor Network (IMN) is a two-year intensive professional learning academy targeted at developing the capacity of Idaho’s Public School Personnel to mentor educators new to the profession.</p> <p>The intent is to develop Mentors who:</p> <ul style="list-style-type: none"> • are a resource for district identified mentees. • use problem solving skills to support the mentee. • are an instructional coach for mentees. • can facilitate opportunities for mentee professional growth. • will collaborate with mentors to improve personal practice and support of mentee. • possess the skills to work with adult learners. • promote a culture of support that included being a trusted listener. <p>Mentors are identified using the following rigorous selection criteria:</p> <ul style="list-style-type: none"> • Recommendation/approval from LEA. • Application process and follow up interview. <p>Recognized in your organization as a change agent, an educator who has credibility among colleagues, and one that is currently a teacher leader or who show great potential as a teacher leader.</p>	<p>Outcomes</p> <p>As a result of the Idaho Mentor Network (IMN):</p> <ul style="list-style-type: none"> • Idaho’s New Teachers will have access to mentors who have both the content knowledge and professional development skills to help education personnel broaden their knowledge base of research-based educational practices. • Students will be engaged in more effective and authentic learning experiences that will result in improved academic achievement with greater success in school and future life experiences. • State, district, and school mentoring policies and procedures will be in place to monitor and support continuous improvement of the instructional core providing strategies, interventions, and resources to all students and education personnel. 	<p>The purpose of the Idaho Mentor Network Project (IMN) is to:</p> <ul style="list-style-type: none"> • help Idaho’s educational system ensure a successful transition from pre-service into the teacher profession. • develop teacher excellence and ensure that every student has an effective teacher. • raise new teacher retention rates and satisfaction . • improve the rigor and consistency of using multiple assessments to guide instruction that is differentiated to meet the diverse learning needs of students. • build norms of collaboration, inquiry, data-driven dialogue and reflection using evidence. • assure the parents and community that new teachers are being supported to attain high levels of professional competence. • ensure that teacher professional development is individualized and based on Professional Teaching Standards and support the Common Core Standards. • develop teacher leadership. • ensure continuous program improvement through ongoing research, development and evaluation.

Delivery of Instruction	Curriculum/Content
<p>Instruction will be delivered via a variety of formats:</p> <p>Face to Face: Participants will enroll in 3- 4 day Mentoring Workshop for graduate credit.</p> <p>Online Graduate Coursework: In addition, Special Education Mentors will enroll in a sequence of four online courses for University Credit that will result in a Consulting Teacher Endorsement from the State of Idaho.</p> <p>Video Conferencing: Participants will also engage in one day videoconferencing events throughout the school year. Mentors will meet at least 5 times using this method. Regional Consultants, Capacity Builders, Idaho Mentor Network staff, and New Teacher Center Staff will be involved in these one day events. IEN origination site will be the training room at SDE. IEN receiving sites will be:</p> <ul style="list-style-type: none"> • BSU • Uofl (Moscow/CDA alternating), • ISU Pocatello/Twin alternating). 	<p>Idaho currently has a contract with the New Teacher Center to deliver 5 Mentor Academies over an 18 month period. The Idaho Department of Education recognizes the Charlotte Danielson’s <i>Framework for Teaching</i> (1996) as an important tool to assess teacher competency, and serve as a model for exemplary teaching. Therefore, Danielson’s <i>Framework for Teaching</i> and the <i>Idaho Core Teaching Standards</i> are both referenced throughout the Mentor Academies</p> <p><i>Year One</i></p> <p><i>Instructional Mentoring & Setting Professional Goals</i> (3 day in person workshop) – June 21 – 23 (SPED Mentors will stay through June 24th), 2011.</p> <p><i>Coaching & Observational Strategies</i> (2 regional session delivered via IEN or Face to Face) September 15 & 16, 2011</p> <p><i>Analysis of Student Work</i> (2 day regional session delivered via IEN) November 17 & 18, 2011.</p> <p><i>Differentiated Instruction</i> (2 day regional session delivered via IEN) March 15 & 16, 2012.</p> <p><i>Year 2</i></p> <p><i>Planning and Designing Professional Development for New Teachers and Mentoring for Equity</i> (4 day in person workshop) – June 19 - 22, 2012</p> <p>Continue coursework for Consulting Teacher Endorsement (see attached)</p>

Responsibilities

SDE	LEA	Mentor	New Teacher Center
<p>Travel, lodging, and per diem costs for participants for Summer Mentor Instructional Leadership Academy (2011). Meal and beverage service for attendees at monthly IEN video conferences.</p> <p>Cost of 3 graduate credits per year (\$916.00).</p> <p>Write for Personnel Improvement Center Grant aimed at recruiting, preparing and retaining special education, early intervention and related services personnel.</p> <p>Provide grant for staff to coordinate Idaho Mentor Network.</p>	<p>Travel costs for participants to attend monthly IEN video conference at their local University.</p> <p>Cost of substitute while mentor is at Academies if needed.</p>	<p>100% attendance at all events.</p> <p>Grade B or better to earn stipend for Graduate Credit.</p> <p>Meet with Mentee at least monthly to practice skills.</p>	<p>Provide curriculum and deliver instruction for Cohort 1 2011- 2012.</p> <p>Provide consultation services to the State of Idaho so that they may begin to develop an sustainable mentor model for 2012 – 2013.</p>

Graduate Certificate, Consulting Teacher Endorsement		
Course Number and Title	Semester Offered	Credits
ED-SPED 551 Tiered Service Delivery Model	Summer 2011	3
ED-SPED 559 Mentoring	Summer 2011	3
Choose one (1) set of three courses from the following:		
ED-SPED 552 Instructional Strategies	Spring 2012	3
ED-SPED 557 Universal Design and Assistive Technology	Fall 2011	3
ED-SPED 558 Data-based Decision Making and Assessment	Fall 2011	3
OR		
ED-SPED 517 School-wide Behavior Support Systems	Spring 2012	3
ED-SPED 518 Intensive, Individualized Behavior Support	Spring 2013	3
ED-SPED 554 Positive Behavior Support	Fall 2012	3
OR		
ED-ECS 511 Early Childhood Special Education Assessment and Evaluation	Fall 2011	3
ED-ECS 514 Early Childhood Special Education Methods	Spring 2012	3
ED-ECS 512 Behavior Support in Early Childhood	Fall 2011	3
OR		
ED-SPED 557 Universal Design and Assistive Technology	Fall 2011	3
Foundations of Secondary Transition	Spring 2012	3
Post-secondary Environments and Interagency Collaboration	Summer 2012	3
TOTAL		15

Consulting Teacher Endorsement (CTE) Core

Tiered Service Delivery Models: Essential components of a responsive instruction and intervention approach, including screening, instruction, intervention, progress monitoring and fidelity of implementation.

Mentoring: Skills and strategies for providing meaningful support and guidance to your fellow teachers, using a variety of coaching styles and mentoring techniques. Develop, implement, and analyze your own coaching plan to lay the foundation for your future as a leader and mentor.

General Special Education Coursework

Universal Design & Assistive Technology: This class will focus on developing an effective core instructional program through the use of Universal Design.

Instructional Strategies: This class focuses on research-based interventions in reading, writing and math to support implementation of Tier 2 activities within an RTI framework.

Data-based Decisions Making & Assessment: Screening, progress-monitoring, academic, behavioral and psychological assessments used to identify students with disabilities and monitor the efficacy of their programs.

Early Childhood Special Education Coursework

Positive Behavior Interventions and Supports in Early Childhood: This class will provide an introduction to positive behavior interventions and supports in early childhood settings with an emphasis on classroom-level implementation. Tier 1 data, systems and practices will be addressed as well as an introduction to Tier 2 and /Tier 2 data and practices.

EI/ECSE Assessments & Evaluation: This class will provide an introduction to assessment and evaluation in early intervention and early childhood special education. The focus will be on screening, eligibility, curriculum-based measurement, progress monitoring, and data-based decision-making.

ECSE Methods: This class will involve the application of a linked system of assessment, goal development, intervention and evaluation to provide services across developmental domains.

Positive Behavior Intervention and Supports (PBIS) Coursework

Positive Behavior Intervention and Supports: This class will provide an introduction to positive behavior interventions and supports in elementary, middle, and high school with an emphasis on classroom-level implementation. Tier 1 data, systems and practices will be addressed as well as an introduction to Tier 2 and /Tier 2 data and practices.

Intensive, Individualized Behavior Support: This class will focus on the data, systems and practices necessary to provide high quality intensive, individualized interventions to students who display chronic problem behavior. Specific content will address functional behavioral assessment and the development of individualized behavior support plans.

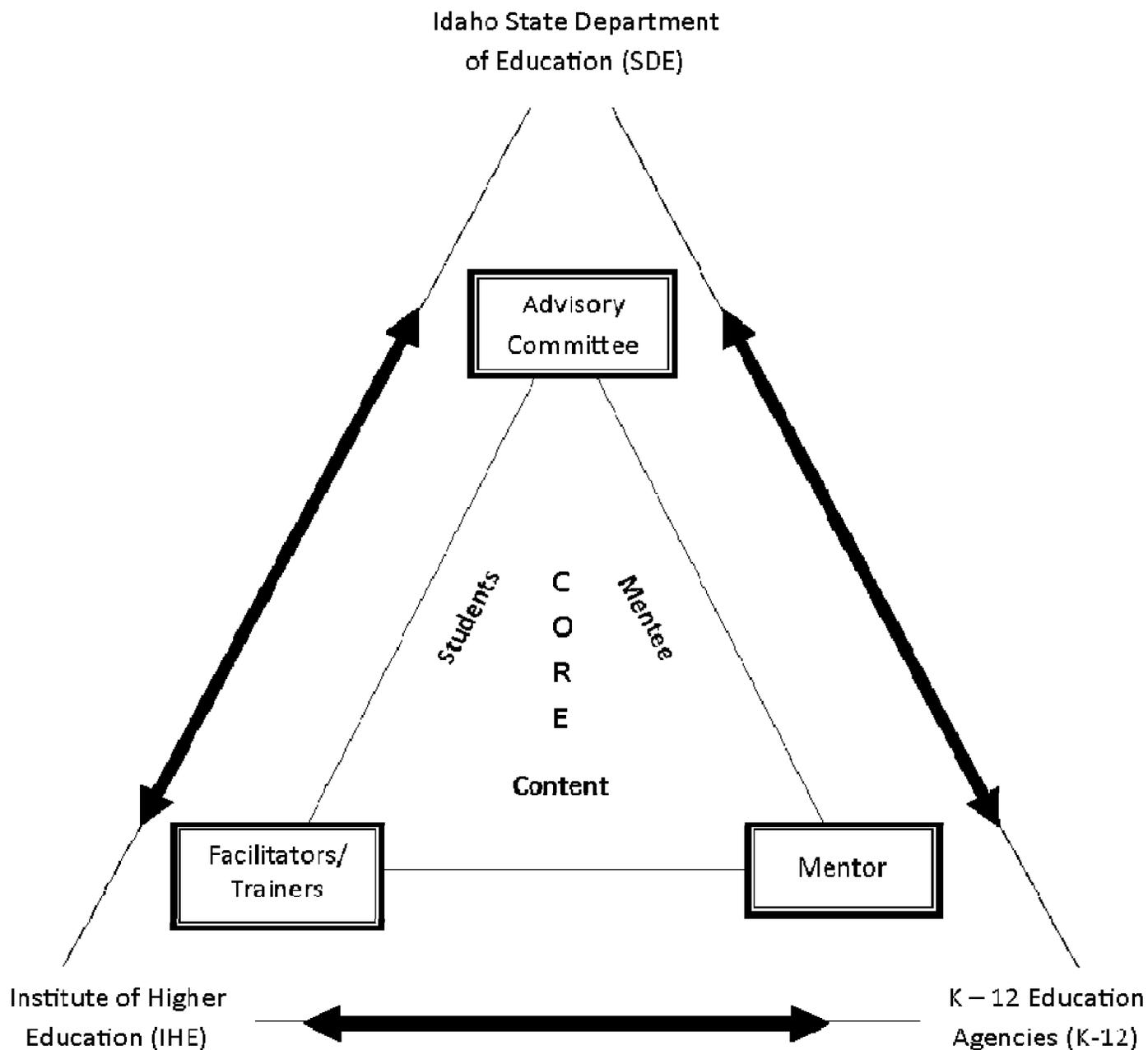
School-wide Behavior Support Systems: This class will focus on school-wide systems of behavior support. Emphasis will be placed on the data, systems and practices necessary across a three-tiered model of behavior support. Students will learn about the readiness requirements, process and considerations for systems-level implementation.

Secondary Transition Coursework

Universal Design & Assistive Technology: This class will focus on developing an effective core instructional program through the use of Universal Design.

Foundations of Secondary Transition: This class will focus on the essential components of career development and transition education for persons with disabilities from middle school through adulthood. Emphasis is placed on IDEA requirements, comprehensive transition assessment, person centered planning, and issues and trends in transition education and services.

Post-secondary Environments and Interagency Collaboration: This class will focus on the skills and strategies for providing meaningful support to transition aged youth with disabilities. Emphasis is placed on Interagency collaboration, post-secondary education supports and services, self-determination, and employment and vocational models.



EXPECTATIONS

Mentee will:

- Self-reflect.
- Describe areas of strengths and weaknesses.
- Ask for help.
- Be open to suggestions to improve instruction.
- Create an environment that welcomes the mentor and fosters an open dialogue for improvement.

Mentor will:

- Become a resource for district identified mentees.
- Use problem solving skills to support the mentee.
- Advocate for the mentee.
- Facilitate opportunities for mentee professional growth.
- Collaborate with mentors to improved personal practice and support of mentee.
- Coach mentees.
- Participate in activities that promote depth of knowledge.
- Promote a culture of support that includes being a trusted listener.

Facilitators/Trainers will:

- Train mentors statewide.
- Model best teaching practices.
- Facilitate the professional learning community among mentors and mentees.

Idaho State Department of Education:

- Provide vision and leadership to support the Idaho Mentor Network Framework.
- Create and implement policies, practices, and procedures that promote the Idaho Mentor Network Framework.
- Dedicate resources to support polices practices and procedures.
- Operationalize Idaho Mentor Network Framework.
- Optimize coordination of services/resources to highest need districts.
- Operationalize statewide evaluation of overall effectiveness.

Program Coordinator will:

- Create rigorous mentor selection process based on qualities of an effective mentor.
- Create and provide ongoing professional development and support for mentors.
- Create a framework that supports a multiyear process.
- Secure funds from the SDE to support process for 3 to 5 year process.
- Collaborate with all stakeholders.
- Research and evaluate program effectiveness.
- Schedule trainings and learning opportunities for Mentors.
- Facilitate the professional learning community among mentors and mentees.

Institute of Higher Education (IHE) will:

- Develop course content to support identified areas of need (RTI, PBIS, ECSE, ST) that can be used by SESTA for professional development content and presentations.
- Deliver courses created for credit (face to face and online).
- Provide input on policy as requested from SDE.
- Research and evaluate program effectiveness as requested.

K-12 Education Agencies (K-12) will:

- Provide administrative support that fosters mentor/mentee participation in the Idaho Mentor Network.
- Create a positive school climate for the support of the program's activities and participate in the ongoing efforts of the Idaho Mentor Network.
- Foster a local network to support the efforts of the Idaho Mentor Network Framework.
- Support mentor/mentee through policies, procedures, practices and incentives that support participants.
- Identify teachers that need support through the use of district evaluations based on the Danielson Framework.

Advisory Committee will: (SESTA, SDE, SSOS, Facilitator, Mentor, Mentee, NTC)

- Provide a platform for stakeholders to provide feedback.
- Meet bi-annually to evaluate program success.

Idaho Mentor Program Standards & Danielson's Framework for Quality Teaching will:

- Provide vision and guidelines for the design and implementation of a high-quality mentor training program for beginning teachers.

Leading the Framework for Teaching

Project: Idaho Mentor Network

Team Members: Jacque Hyatt, Becky Martin, Teresa Burgess, NTC, Christina Linder, Carol Carnahan

Funding Source: SPDG Grant

Purpose: The Idaho Mentor Network (IMN) supports the development and implementation of quality mentor programs in Idaho that mentor educators new to the profession.

Method of Delivery: 5 2 day mentor academies delivered over 18 months in a face to face format utilizing the New Teacher Center program and staff.

Action Steps What Will Be Done?	Responsibilities Who Will Do It?	Timeline By When? (Day/Month)	Resources/Barriers A. Resources available/needed B. Barriers present/perceived	Communications Plan Who is involved? What methods? How often?
Step 1: Confirm time for NTC sharing at the IHE Meeting on February 1st	Christina	January 15th	Waiting for schedule change for other agenda items means we won't know if we have a morning slot or an afternoon slot.	Christina will email info to team members as soon as it is available so Becky and Jacque can schedule some planning time with NTC staff
Step 2: Idaho IHE Partnership Meeting February 9th and 10th <ul style="list-style-type: none"> • Set agenda • Schedule room (Barbara Morgan) • Identify districts • Invite Districts 	Agenda set ☺ 9 th IHE only morning of 10 th IHE & K12 afternoon of 10 th is K12 Room booked ☺ Compile list of districts and personnel who have participated in	December 22 nd ☺	Agenda for February 9 th and 10 th List of districts to invite to Feb 9 th & 10 th Letter to districts inviting them to 9 th	Katie set a tentative agenda, Katie booked room Teresa and Joe compiled list and emailed to team on December 22 nd ☺

	<p>the IMN since 2006.</p> <p>Review list and Invite (Jacque and Christina)</p>	<p>January 3rd</p>	<p>and 10th meeting</p>	<p>Email districts invited and letter to team when completed.</p>
<p>Step 3: How do we assess what mentors in the current cohort are doing? How do we get mentors in our current cohort mentoring? Define instructional coach and mentor Look at district systems across the state Look at School Improvement plans</p>	<p>Jacque, Becky, Teresa, and Carol will meet to determine how these questions can be answered and contract will be put in place for Carol to assist.</p>			
<p>Step 4:</p>			<p>A.</p> <p>B.</p>	
<p>Step 5:</p>			<p>A.</p> <p>B.</p>	

Project: Danielson for Professional Practice Project
Team Members: Jacque, Becky, Christina, Teresa, Joanie, Kathleen, Carol

Purpose: Jacque needs to define purpose of this project in Jacquees ☺

Action Steps What Will Be Done?	Responsibilities Who Will Do It?	Timeline By When? (Day/Month)	Resources/Barriers C. Resources available/needed D. Barriers present/perceived	Communications Plan Who is involved? What methods? How often?
Step 1: Develop 4 day Peer Coaching Training	Kathleen, Joanie, Carol, Jacque	December 22 nd ☺	Delivery of training. How, when, where? All to be determined after IHE and K12 partnership meetings in February. Possible Summer Institute 3 days in June 1 follow up	Carol and Jacque meet and worked with Kathleen and Joanie to develop outline for 4 day training and Kathleen and Joanie will deliver training binders in December of 2011.
Step 2: Schedule IMN Meeting for March 9 th in the afternoon. Set delivery schedule for Peer Coaching	Jacque	March 14 th		Jacque will email training dates after our IMN on March 15th
Step 3: Summer eMSS training	Carol and Becky			

Step 4:			A. B.	
Step 5:			A. B.	

Project: Danielson for Evaluation Project
Team Members: Becky Martin, Christina Linder, Teresa Burgess, Kathleen Hanson, Joanie Peterson, Rob Sauer
Funding Source: Title IIA
Purpose: Alignment to Danielson and promotion of Interrater Reliability and Fidelity throughout the state for teacher evaluations.
Target Audience: Administrator, principals, evaluators, teacher leaders
Objective: To provide statewide trainings for all evaluators concerning proficiency assessment for Danielson Framework.
Method of Delivery and outcome:
Phase 1: 2011 -12 & 2012 2013(Fall) Statewide face to face 4 day regional trainings delivered by Joanie Peterson and Kathleen Hanson
 Phase II: 2012-2013 – Spring - Online followup
 Phase III: Proficiency Exam – Statutory Change??? – Change Board Rule? Implement for recertification 2015 – 2020 – School Board push??

Hope is to have Recertification by 2015-2020

Action Steps What Will Be Done?	Responsibilities Who Will Do It?	Timeline By When? (Day/Month)	Tasks and Resources/Barriers E. Resources available/needed F. Barriers present/perceived	Communications Plan Who is involved? What methods? How often?
Step 1: Develop 4 day training	Katheleen Hanson and Joanie Peterson	12/19/2011		Completed
Step 2: <ul style="list-style-type: none"> • Contracts for trainer • Schedule training 	Christina	January 17th	Schedule trainings: Region 3: January 18th, March 8th, April 19th, June 14th	Teresa, Becky, Christina

<ul style="list-style-type: none"> • Invite districts 			<p>Pocatello: January 20th, March 6th, April 24th June 7th</p>	
<p>Step 3: Preassessment of each district attending training for day 1 – Teachscape Proficiency Online preassessment</p>	<p>Becky</p>		<p>Contact Joanie and Kathleen to see if any survey was done for CDA Content vs. practice</p>	<p>Christina, Teresa, Becky will take the pre-test to see how they can use that with this training to collect data.</p>
<p>Step 4: Create an evaluation for training</p>		<p>Completed</p>	<p>Locate evaluation and review for data points</p>	
<p>Step 5: Measure Impact and Write Project Report</p>	<p>Becky & Chrstina</p>			
<p>Step 6: Plan for Phase II and Phase III</p>	<p>Team</p>		<ul style="list-style-type: none"> ○ Basic Danielson Framework and observation and testing interrater reliability • Districts should come knowing the basics of Danielson – content knowledge – make available online – book study • How can we do pre-assessments so we can differentiate the instruction and build choice and buy-in? • How can we deliver the Basic Danielson Training and Observation Training online? ISEE & IEN???? • Administrator Evaluation Focus Groups – Show Teachscape 	<p>Team meetings quarterly</p>

			capabilities Phase III Proficiency Exam – Statutory Change??? – Change Board Rule? Implement for recertification 2015 – 2020 – School Board push??	
			How many years have you been using Danielson? Multiple measures	



Alternative Measures of Teacher Performance

The American Race to the Top and the Race to the Top grant pushed states and districts in the development of high-quality teacher systems that focus on student growth as a measure of teacher effectiveness and use multiple measures to inform critical decisions relating to opportunity, improvement and career advancement.

Policy-to-Practice Brief

Alternative Measures of Teacher Performance

May 2011

Amy Potemski

Mahua Baral

Cassandra Meyer

With Lisa S. Johnson, Ed.D.
and Sabrina W. M. Laine, Ph.D.



CONTENTS

The Changing Policy Landscape 1

The Importance of Alternative Measures of Teacher Performance 2

 Defining Effectiveness..... 2

Examples of Alternative Measures 4

Research on Alternative Measures 5

Conclusion 6

References 7

Appendixes

 Appendix A. Gallup Student Poll 13

 Appendix B. Scoop Notebook: Examining Classroom Artifacts 15

 Appendix C. Surveys of Enacted Curriculum 17

 Appendix D. Teacher Portfolios..... 19

 Appendix E. Tripod Surveys: Student, Teacher, and Parent Surveys..... 22

THE CHANGING POLICY LANDSCAPE

The American Recovery and Reinvestment Act (ARRA) of 2009 funneled an unprecedented amount of federal funding to education initiatives through a variety of funding streams. By now, most education stakeholders are aware of the four primary assurances outlined in ARRA and made available to states through the Race to the Top competitive grant:¹

- “Adopting standards and assessments that prepare students to succeed in college and the workplace to compete in the global economy.”
- “Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction.”
- “Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most.”
- “Turning around our lowest achieving schools” (U.S. Department of Education, 2009, p. 2).

Since the passage of ARRA, these assurances have driven changes in state legislation, especially as states prepared to participate in the Race to the Top competitive grant program. In a review of the 41 applications submitted for Phase I of Race to the Top, Learning Point Associates (2010b), an affiliate of American Institutes for Research, found that 29 (71 percent) of the 41 applications submitted by states and the District of Columbia included descriptions of recently passed legislation or intentions to introduce legislation in support of Race to the Top program priorities. Specific to teacher evaluation, a total of 11 states passed, or expressed an intention to pass, legislation related to teacher evaluation in the following key areas: prescribing measures to evaluate

teachers (7 states), prescribing the use of evaluation data (2 states), and prescribing both measures to evaluate teachers and the use of evaluation data (2 states).

Although most new state laws focused on the use of student achievement data to assess teacher performance, another common theme in the legislation was the redesign of educator evaluation systems at the state and district levels, including the stated use of observation rubrics and other measures of teacher performance (Learning Point Associates, 2010b).

In addition to ARRA, the Common Core State Standards movement, spearheaded by the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO), provides states with an additional incentive to agree on definitions for the essential knowledge and skills necessary to the future success of K–12 students. NGA and CCSSO worked collaboratively with states, educators, content experts, researchers, national organizations, and community groups to ensure that stakeholders had a significant role in the development process. Forty-one states, the District of Columbia, and the U.S. Virgin Islands have adopted the Common Core State Standards.² Currently, state standards are available in mathematics and English language arts, which also include literacy in history/social studies, science, and technical subjects. NGA and CCSSO also consider the application of the standards to English learners and students with disabilities.

This Policy-to-Practice Brief introduces five current examples of measures of teacher performance. The goal is to assist regional comprehensive centers and state education agencies in building local capacity to incorporate the use of alternative measures of teacher performance into the overhaul of state evaluation systems—especially in states with looming legislative deadlines.

¹ For a complete listing of education programs under ARRA as well as links to regulations, guidance, and resources provided by the U.S. Department of Education, visit

² For more information on the states and territories that have adopted the Common Core State Standards as well as links to the detailed standards, guidance, and other resources, visit

THE IMPORTANCE OF ALTERNATIVE MEASURES OF TEACHER PERFORMANCE

ARRA and the Race to the Top grant program have pushed states and districts to invest in the development of high-quality teacher evaluation systems. Such systems have two specific elements:

- A focus on student growth data as a measure of teacher effectiveness
- Multiple measures to inform critical decisions relating to opportunities for teacher improvement and career advancement (e.g., promotion, tenure, equitable distribution, compensation).

Historically, most states and districts have used classroom observations as the primary tool to assess teacher performance (Brandt, Thomas, & Burke, 2008; Weisberg, Sexton, Mulhern, & Keeling, 2009). Although classroom observations—in combination with student growth measures—provide multiple data points on teacher performance, additional alternative measures also should be considered to ensure a robust teacher evaluation system that captures the many facets of effective teaching.

Alternative measures can take many forms, from student engagement surveys to teacher portfolios. It is beyond the scope of this brief to cover every alternative measure to assess teacher effectiveness; however, the brief highlights five measures that are included in the online *Guide to Teacher Evaluation Products* (National Comprehensive Center for Teacher Quality, 2010) as examples of alternative measures that have potential for use in teacher evaluation.

In a review of teacher evaluation reforms proposed in state Phase 1 Race to the Top applications, Learning Point Associates (2010a) found that in addition to student growth measures, states also discussed plans to develop multiple measures of teacher performance beyond student learning. Although most state applications included references to observation rubrics, some states also described other measures of teacher performance in their applications, including the following (Learning Point Associates, 2010a):

- A review of classroom artifacts or portfolios submitted by the teacher
- Teacher planning, instructional, and assessment artifacts (6 states)
- Teacher self-reflection portfolios (5 states)
- Examples of student work (3 states)
- Provisions for peer review and feedback (6 states)
- Student reflections and feedback (5 states)
- Teacher participation in professional development (1 state)
- Follow-up work on teacher adaptation of classroom practices in response to feedback from formal and informal observations (1 state).

Defining Effectiveness

Understanding that student growth measures on their own have limitations for determining “effective” and “highly effective” designations for teachers and leaders, the U.S. Department of Education (2009) has reinforced the need to include multiple measures of teacher performance as the most robust approach to fully capturing classroom practice (See “Definitions of Effective and Highly Effective Teachers”).


DEFINITIONS OF EFFECTIVE AND HIGHLY EFFECTIVE TEACHERS

The U.S. Department of Education (2009, p. 12) provides the following definitions of effective and highly effective teachers:

Effective teacher means a teacher whose students achieve acceptable rates (e.g., at least one grade level in an academic year) of student growth (as defined in this notice). States, LEAs [local education agencies], or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, multiple observation-based assessments of teacher performance.

Highly effective teacher means a teacher whose students achieve high rates (e.g., one and one-half grade levels in an academic year) of student growth (as defined in this notice). States, LEAs, or schools must include multiple measures, provided that teacher effectiveness is evaluated, in significant part, by student growth (as defined in this notice). Supplemental measures may include, for example, multiple observation-based assessments of teacher performance or evidence of leadership roles (which may include mentoring or leading professional learning communities) that increase the effectiveness of other teachers in the school or LEA.

In its 2008 review of existing research on evaluating teacher effectiveness, the TQ Center introduced a five-point definition of *teacher effectiveness* that was intended to initiate state and regional conversations on the types of measures that might be needed to determine effective classroom teaching (Goe, Bell, & Little, 2008). The TQ Center's definition recognizes the primacy of student growth data, but it also highlights additional important aspects of teaching, many of which are not currently measured through teacher observations or student learning growth measures. This definition highlights a specific need for alternative measures of teacher performance to determine effectiveness.

Given the significant policy focus on reforming state and local teacher evaluation systems that include multiple measures of teacher performance, there is a clear need for the following:

- The development of products and services that provide alternative measures of teacher performance
- Widespread dissemination of the products and services for states to respond to legislative initiatives implemented since the passage of ARRA

“The five-point definition of *teacher effectiveness* consists of the following:

- Effective teachers have high expectations for all students and help students learn, as measured by value-added or other test-based growth measures, or by alternative measures.
- Effective teachers contribute to positive academic, attitudinal, and social outcomes for students such as regular attendance, on-time promotion to the next grade, on-time graduation, self-efficacy, and cooperative behavior.
- Effective teachers use diverse resources to plan and structure engaging learning opportunities; monitor student progress formatively, adapting instruction as needed; and evaluate learning using multiple sources of evidence.
- Effective teachers contribute to the development of classrooms and schools that value diversity and civic-mindedness.
- Effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure student success, particularly the success of students with special needs and those at high risk for failure.” (Goe et al., 2008, p. 8)

EXAMPLES OF ALTERNATIVE MEASURES

States and districts recently have begun to implement teacher evaluation reforms. Table 1 provides information on five alternative measures of teacher performance that might be used to supplement growth measures and observation rubrics. (For additional information about these measures, refer to Appendixes A–E.)

Table 1. Five Alternative Measures of Teacher Performance

Alternative Measure Product or Service or Service	Developer	Type of Information Gathered	Cost of Product
Gallup Student Poll	Gallup Inc. America’s Promise Alliance American Association of School Administrators	The poll is administered to students in Grades 5-12. The poll measures three variables identified as key factors that drive students’ grades: hope, engagement, and well-being.	Registered public schools and districts can use this measure at no cost.
Scoop Notebook	National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at the Center for the Study of Evaluation (CSE) RAND Corporation Stanford University	This measure uses artifacts and related materials to represent classroom practice. Artifacts and other materials can include the following: lesson handouts; student classwork; homework; photos of classroom layout, equipment, and board work; teacher reflections on each lesson.	States may use publically available research and resources to implement this measure in their schools at no cost. To receive expert assistance to use the tool, states may negotiate pricing with the developers.
Surveys of Enacted Curriculum (SEC)	Council of Chief State School Officers (CCSSO) Wisconsin Center for Education Research (WCER)	Teachers report information on subject coverage, length of time spent on topics, and cognitive depth covered in their classroom instruction through an online survey. Teachers as well as school, district, and state leaders can use this information to inform professional development and assess the extent to which teacher instruction aligns with state standards and assessments.	Cost for tools and services varies and is determined by CCSSO and WCER on a case-by-case basis.
Teacher Portfolios	Varies, based on specific example (See Appendix D.)	Teachers pull together portfolios that can include the following: <ul style="list-style-type: none"> • Video clips • Lesson plans • Teacher self-assessments or evaluations • Examples of student work 	Costs vary, depending on whether portfolios are developed in-house or with consultant. (See Appendix D for more details.)

Alternative Measure Product or Service or Service	Developer	Type of Information Gathered	Cost of Product
Tripod Surveys	Harvard University	<p>This measure consists of surveys for students, teachers, and parents.</p> <p>The surveys identify attitudes, perceptions, experiences, and classroom practice related to teacher content knowledge, pedagogical knowledge, and relationships between teachers and students.</p> <p>The surveys examine the Seven C's of teacher quality:</p> <ul style="list-style-type: none"> • Care about students • Control of student behavior • Captivating students • Clarifying lessons • Challenging students academically • Conferring with students • Consolidating knowledge 	<p>Costs vary, and consultation services are customized based on client needs.</p>

For additional information, please refer to the online *Guide to Teacher Evaluation Products* and Appendixes A-E.

The products and services included in Table 1 align with the types of alternative measures specified in some of the state Race to the Top applications, such as review of classroom artifacts or portfolios; teacher planning, instructional, and assessment artifacts; teacher self-reflection portfolios; examples of student work; provisions for peer review and feedback; and student reflections and feedback (Learning Point Associates, 2010a). The products and services were selected from the range of products available in more detail in the TQ Center's online *Guide to Teacher Evaluation Products* (2010).

RESEARCH ON ALTERNATIVE MEASURES

Although further evaluation and research is needed to fully understand the best way to fit these measures into teacher evaluation practices, Table 2 provides a short synopsis of the advantages and challenges identified by currently available research.

Table 2. Advantages and Challenges of Alternative Measures

Measure of Teacher Performance	Research Cited*	Advantages	Challenges
Gallup Student Poll	America's Promise Alliance (2010) Gallup Consulting Education Practice (2009) Lopez (2010) Lopez, Agrawal, and Calderon (2010)	The poll is available through a secure, online administration website. Students can complete the poll in less than 10 minutes. For a fee, Gallup provides analysis of the data that correlate survey results with grade-level or classroom-level gains.	The poll is not an alternative measure for all students, as it is not available before Grade 5. The poll requires Internet access.
Scoop Notebook	Borko, Stecher, Alonzo, Moncure, and McClam (2005) Borko, Stecher, and Kuffner (2007) Stecher et al. (2005)	This measure can increase teacher commitment to the evaluation process. Schools and districts may be able to better address the professional development needs of teachers with the critical information gleaned from this measure. This measure may assist teachers in analyzing student work in professional learning communities.	Only mathematical and science rating guides are currently available. It might be difficult to develop as a rigorous and comparable measure of teacher effectiveness as part of a high-stakes evaluation system. It may not be useful as a measure in classrooms that produce minimal artifacts (e.g., physical education). This approach takes time and effort to complete.
Surveys of Enacted Curriculum	Blank (2004) Blank, Porter, and Smithson (2001) Council of Chief State School Officers (2004) Council of Chief State School Officers (2010)	The SEC collect a large amount of information on teacher practice. The SEC report on instructional practice across a school year, which can be difficult information to obtain through other types of evaluation measures.	This measure relies on teacher self-reporting, which may not be accurate. This measure requires training for teachers and administrators to view and understand the data to be used most effectively.

Measure of Teacher Performance	Research Cited*	Advantages	Challenges
Teacher Portfolios	Goe, Bell, and Little (2008) Little, Goe, and Bell (2009) National Comprehensive Center for Teacher Quality (2010)	Teachers collect and reflect on evidence across various activities, which encourages a perspective on teaching beyond the classroom. If conducted collaboratively, this measure can create a more cohesive teaching team. Receiving and providing support to colleagues may promote professional growth. This measure can be conducted in an online format or through a physical collection of artifacts.	Feedback is time-sensitive. It is best to apply this measure over the course of a year; however, it is difficult to regulate. There is tension between using evidence as part of an evaluation or for professional growth.
Tripod Surveys	Bill & Melinda Gates Foundation (2010) Ferguson (2002a) Ferguson (2002b)	This measure can be used to report otherwise unobservable factors that may affect teaching, such as knowledge, intentions, expectations, and beliefs. The surveys provide the unique perspective of the teacher as well as the perspective of students, who have the greatest amount of experience with teachers. This measure can provide formative information to help teachers improve practice in a way that connects with students. This measure makes use of the perspective of students who may be as capable as adult raters of providing accurate ratings.	This measure relies on teacher self-reporting, which may not be accurate. Students cannot provide information on certain aspects of teaching, such as a teacher's content knowledge, curriculum fulfillment, or professional activities.

*For full references, see Appendixes A-E.

As evidenced in Table 2, each measure has distinct advantages and implementation challenges. In some cases, such as the Gallup Student Poll and the Tripod Surveys, the relatively small cost of implementation is advantageous. However, it is also important to take into account the state's or district's specific teacher evaluation needs.

CONCLUSION

As state and district efforts continue to focus on teacher evaluation system reform, it is necessary to explore options for the gradual inclusion of multiple measures of performance to accurately evaluate teacher effectiveness. As state and district staff consider the five alternative measures presented in this brief, they should reflect on the following questions:

- What teaching standards is the system trying to measure?
- What kind of support can the state provide to LEAs for implementation?
- How will the evaluation system be used?
 - Guiding professional development
 - Certification or tenure decisions
 - Teacher career ladders
 - Alternative compensation programs
 - Addressing the inequitable distribution of teachers

For a more in-depth look at making decisions regarding state and district teacher evaluation systems, see the *Practical Guide to Designing Comprehensive Teacher Evaluation Systems*

. This guide walks states and districts through questions that are essential to the development and implementation of a high-quality, comprehensive teacher evaluation system.

The advantages and implementation challenges of the alternative measures presented in this brief directly relate to the type of outcomes affected by the evaluation system. States and districts should carefully review examples of each measure in practice and determine the appropriate measures in the context of their school systems.

REFERENCES

- American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115 (2009). Retrieved May 6, 2011, from
- Brandt, C., Thomas, J., & Burke, M. (2008). *State policies on teacher evaluation practices in the Midwest region* (REL Technical Brief, REL 2008-004). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved May 6, 2011, from
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved May 6, 2011, from
- Learning Point Associates. (2010a). *Evaluating teacher effectiveness: Emerging trends reflected in the state phase I Race to the Top applications*. Naperville, IL: Author. Retrieved May 6, 2011, from
- Learning Point Associates. (2010b). *State legislation: Emerging trends reflected in the state phase I Race to the Top applications*. Naperville, IL: Author. Retrieved May 6, 2011, from
- National Comprehensive Center for Teacher Quality. (2010). *Guide to teacher evaluation products* [Website]. Washington, DC: Author. Retrieved May 6, 2011, from
- U.S. Department of Education. (2009). *Race to the Top program: Executive summary*. Washington, DC: Author. Retrieved May 6, 2011, from
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project. Retrieved May 6, 2011, from

APPENDIXES

APPENDIX A. GALLUP STUDENT POLL

Developer of Product and Services

The Gallup Student Poll was designed by Gallup Inc., in partnership with America's Promise Alliance and the American Association of School Administrators.

Description of Product and Services Available

In 2009, Gallup Inc. launched the Gallup Student Poll, a school-based online survey for students in Grades 5–12 that measures three variables: hope, engagement, and well-being. Gallup Inc. defines *hope* as “the ideas and energy students have for the future,” *engagement* as a student’s “level of involvement in and enthusiasm for school,” and *well-being* as “how students think about and experience their lives” (see America’s Promise Alliance, 2010, listed in the Research and Resources section at the end of this appendix). Through extensive research, these three variables were identified as key factors that drive students’ grades, achievement scores, retention, and future employment. Furthermore, research has revealed that the variables are linked to teacher talent and teacher engagement; staff and student engagement have been shown to drive positive outcomes and explain variance in school performance (see Gallup Consulting Education Practice, 2009, listed in the Research and Resources section).

The survey is administered once during each school year. Students can access the survey on a secure website using a registered account. The survey takes, on average, less than 10 minutes to complete. In addition to several demographic questions (e.g., age, grade, gender), students are asked 20 core questions about their perspectives related to their home, school, and community lives. Survey questions were first developed in 2006 and have since been reviewed and refined based on additional research, focus group feedback, and psychometric studies conducted from 2008 to 2010. Studies include a 2008 expert review of items, pilot studies in 2008 and 2009, representative panel studies in 2009 and 2010, and a 2009 validation study.

In 2009 and 2010, more than 450,000 students from across the country took the survey. Data from the survey have been used by schools and districts to build student and staff engagement and to provide information on how to select strategic initiatives, trainings, and interventions.

Training for Use of Product and Services

Gallup Inc. has developed a webinar series to communicate information about the Gallup Student Poll to educators and community leaders. The webinars are free and are offered throughout the year. For a schedule of upcoming webinars, please visit the *Online Learning & Webinars* webpage

Cost of Product and Services

The survey is free for registered public schools and districts.

Advantages and Implementation Challenges

Advantages

- Free of charge.
- Available online through a secure website.
- Takes less than 10 minutes to complete.

Implementation Challenges

- Not available for students prior to Grade 5.
- Requires computers with Internet access.

How States Can Get More Information

States can get more information at the Gallup Student Poll website .

Technical support, provided by Gallup Inc. is available by phone (866-346-4408) Monday through Thursday from 7 a.m. to 7 p.m. and Friday from 7 a.m. to 5 p.m. (Central Time).

Research and Resources

America's Promise Alliance. (2010, August 12). *Gallup student poll finds gap between perception and reality in youth hope, engagement and wellbeing* [Press release]. Retrieved May 6, 2011, from

Gallup Consulting Education Practice. (2009). *Building engaged schools: A scientific method for improving school performance* [Brochure]. Washington, DC: Author. Retrieved May 6, 2011, from

Lopez, S. J. (2010). *Youth readiness for the future: A report on findings from a representative Gallup Student Poll sample*. Washington, DC: Gallup Inc. Retrieved May 6, 2011, from

Lopez, S. J., Agrawal, S., & Calderon, V. J. (2010). *The Gallup Student Poll technical report*. Washington, DC: Gallup Inc. Retrieved May 6, 2011, from

APPENDIX B. SCOOP NOTEBOOK: EXAMINING CLASSROOM ARTIFACTS

Developer of Product and Services

The Scoop Notebook was developed by the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at the Center for the Study of Evaluation (CSE); RAND Corporation; and Stanford University.

Description of the Product and Services Available

The Scoop Notebook is a protocol for gathering and rating the quality of middle school mathematics and science classroom artifacts. It was developed through a five-year project funded through CRESST. The goal of the project was to use artifacts and related materials to represent classroom practice well enough that a person unfamiliar with a teacher or lessons can make valid judgments about selected features of practice solely on the basis of those materials. Moreover, there are two potential uses of the Scoop Notebook: as part of a system of multiple measures to characterize teacher effectiveness or as a formative tool for teacher professional development.

During the course of one week, teachers collect artifacts and other materials (e.g., lesson handouts; student classwork; homework; photos of classroom layout, equipment, and board work; teacher reflections on each lesson) and put them in a binder called the “Scoop Notebook.” (Articles and studies listed in the Research and Resources section at the end of this appendix provide detailed instructions on creating the binders and using rubrics to analyze artifacts.) Rating guides for the notebook are based on previous research, the National Science Education Standards, and Principles and Standards for School Mathematics. Although the tool was developed and field-tested in middle school classrooms, the developers believe it is appropriate for other grade levels as well.

During the five-year project, developers have conducted numerous studies to develop, refine, and test the reliability and validity of the product (see the Research and Resources section at the end of this appendix). Between 2003 and 2007, the Scoop Notebook was tested and used successfully in 36 middle schools in Los Angeles and Denver. Because the tool is publicly available, it may be used in multiple schools and districts beyond the developers’ knowledge.

Training for Use of Product and Services

The Scoop Notebook can be used without training. All materials and guidebooks are available online at no cost to the user. Questions concerning the specific use of the notebook can be addressed to the developers listed in the How States Can Get More Information section.

Cost of Product and Services

States may use publically available research and resources (see the Research and Resources section) to implement the Scoop Notebook in their schools, free of cost. To receive expert assistance to use the tool, states may negotiate pricing with the developers (see the How States Can Get More Information section).

Advantages and Implementation Challenges

Advantages

- Free online; additional expert assistance available for a fee.
- May increase teacher commitment to the evaluation process.
- May provide schools and districts critical information to better address professional development needs of teachers.
- May assist teachers in analyzing student work in professional learning committees.

Implementation Challenges

- Currently, only mathematics and science ratings guides available.
- May be difficult to develop as a rigorous and comparable measure of teacher effectiveness.
- May not be useful for a measurement of classrooms that produce minimal artifacts (e.g., physical education).
- Takes time and effort to complete.

How States Can Get More Information

States can get more information from the developers of the product:

- Dr. Hilda Borko (650-723-7640, hildab@stanford.edu)
- Dr. Brian Stecher (310-393-0411, brian_stecher@rand.org)

Research and Resources

Borko, H., Stecher, B. M., Alonzo, A. C., Moncure, S., & McClam, S. (2005). Artifact packages for characterizing classroom practice: A pilot study. *Educational Assessment, 10*(2), 73–104.

Borko, H., Stecher, B., & Kuffner, K. (2007). *Using artifacts to characterize reform-oriented instruction: The Scoop Notebook and rating guide* (CSE Technical Report No. 707). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (ERIC Document Reproduction Service No. ED495853). Retrieved May 6, 2011, from

Stecher, B., Wood, A. C., Gilbert, M. L., Borko, H., Kuffner, K. L., Arnold, S. C., et al. (2005). *Using classroom artifacts to measure instructional practices in middle school mathematics: A two-state field test* (CSE Report No. 662). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved May 6, 2011, from

APPENDIX C. SURVEYS OF ENACTED CURRICULUM

Developer of Product and Services

The *Surveys of Enacted Curriculum* (SEC) were developed by the Council of Chief State School Officers (CCSSO) and the Wisconsin Center for Education Research (WCER).

Description of Product and Services Available

The SEC are online surveys that ask teachers to report information on subject coverage, length of time spent on topics, and cognitive depth covered in their classroom instruction. Teacher results can be compared with the content included in state standards and state assessments.

Using aggregated information from several teachers, administrators at the school, district, and state levels can identify the extent to which teacher instruction aligns with state standards and state assessments and use this information to inform professional development and school improvement. By tracking this information over time, the SEC can provide feedback to schools, districts, and states on program implementation.

Individual teachers also can review their practice and compare it with standards and the results of other teachers in their school or district. Consequently, it is possible for SEC data to be part of the information that teachers consider when self-evaluating their performance.

This tool was designed for Grades K–12 mathematics, science, and language arts teachers. Mathematics and science surveys were written and field-tested from 1994 to 1998, with English language arts surveys and reports developed from 2002 to 2003. Eleven states are part of the SEC State Collaborative on Assessment and Student Standards: Arkansas, Idaho, Indiana, Iowa, Kansas, Mississippi, Missouri, Montana, Ohio, Oregon, and Wisconsin.

The final report of the SEC, a study of the mathematics and science measures across 11 states, was funded through a grant by the National Science Foundation and published in 2001. It includes information on measure validity and ways to mitigate issues related to teacher self-reporting on practice (see Blank, Porter, & Smithson, 2001, listed in the Research and Resources section at the end of this appendix).

The Common Core State Standards recently were analyzed for their content, and the results are publically available (see Council of Chief State School Officers, 2010, listed in the Research and Resources section). Several states are currently using the SEC to consider the alignment between instruction and the Common Core State Standards. Currently, SEC instruments are being adapted and expanded to facilitate a deeper examination of the instruction that students with disabilities receive. In addition, there are plans to develop a teacher-log format as well as a tool that would allow teachers to study the *intended* curriculum as compared with the *enacted* and *assessed* curriculum.

Training for Use of Product and Services

Training can be scheduled by contacting CCSSO or WCER. Resources related to training can be found at the SEC Resources webpage

Cost of Product and Services

Costs of tools and services vary and can be determined by contacting CCSSO or WCER.

Advantages and Implementation Challenges

Advantages

- Collects a large amount of information on teacher practice.
- Reports on instructional practice across a school year—information that is difficult to obtain through other types of evaluation measures.

Implementation Challenges

- Relies on teacher self-reporting, which may not be accurate.
- Requires training for teachers and administrators to view and understand the data so they may be used most effectively.

How States Can Get More Information

States can find more information at the CCSSO SEC webpage and the WCER SEC webpage or by contacting the following:

- Rolf K. Blank (202-336-7044; rolfb@ccsso.org)
- John Smithson (608-263-4354; johns@education.wisc.edu)

Research and Resources

Blank, R. K. (2004, April). *Findings on alignment of instruction using enacted curriculum data: Results from urban schools*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA. Retrieved May 6, 2011, from

Blank, R. K., Porter, A., & Smithson, J. (2001). *New tools for analyzing teaching, curriculum and standards in mathematics and science: Results from Survey of Enacted Curriculum Project final report*. Washington, DC: Council of Chief State School Officers. Retrieved May 6, 2011, from

Council of Chief State School Officers. (2004). *Data on enacted curriculum study: Summary of findings*. Washington, DC: Author.

Council of Chief State School Officers. (2010, September 20). *Content analysis of Common Core State Standards: Initial findings* [PowerPoint presentation]. Retrieved May 6, 2011, from

APPENDIX D. TEACHER PORTFOLIOS

Developer of the Product and Services

Teacher portfolios have been developed by various state education agencies, local education agencies, and education organizations.

Description of the Product and Services Available

Following are some examples of teaching portfolios.

Washington ProTeach Portfolio

The ProTeach portfolio collects the following student-based evidence to measure teacher effectiveness:

- Professional growth and contributions. Includes analysis and reflection on professional growth and its impact on student learning.
- Building a learning community. Includes a description and analysis of the learning environment established in the single class or classroom.
- Curriculum, instruction, and assessment. Includes analysis and reflection of the curriculum, instruction, and assessment and their impact on three focus students.

Artifacts collected for the portfolio include teacher and student work, written commentary, and samples in student voice (e.g., evidence of student learning from the students' perspective).

Alexandria (Virginia) City Public Schools—Performance Evaluation Program

The Performance Evaluation Program has four components: formal observations, informal observations, teacher portfolios, and academic goal-setting. The teacher portfolios are made up of artifacts that provide documents for 17 performance responsibilities, determined by Alexandria City Public Schools.

Performance Assessment for California Teachers—Teaching Event

Teaching Event is a teacher portfolio modeled after the teacher portfolio assessments of the Connecticut Department of Education, Interstate Teacher Assessment and Support Consortium, and National Board for Professional Teaching Standards. It documents work that meets criteria for six components: context, planning, instruction, assessment, reflection, and academic language. The goal is to have teacher candidates make connections between the different tasks and to provide evidence from a brief learning segment in depth. The directions for constructing the Teaching Event portfolio are designed to direct teacher candidates to plan, teach, and reflect on their teaching within the specific context of their students and their learning. Teaching Event portfolios include video clips, scorers with subject-specific expertise, and subject-specific benchmarks. Training is provided on its use.

National Board for Professional Teaching Standards—National Board Certification

National Board Certification is a standards-based assessment of teacher effectiveness. A score reflects the degree to which assessors were able to locate clear, consistent, and convincing evidence that the candidate has met the standards specific to his or her certificate field. The National Board Certification process consists of a teacher portfolio as well as other components. The portfolios are required to contain four entries. Three of these entries are classroom based; the fourth requires working with families and the larger community and with colleagues and the larger profession. At least two of the classroom-based entries must use video recording. In addition, teachers must provide a collection of student work as well as commentary describing, analyzing, and reflecting on the evidence.

The National Board for Professional Teaching Standards has conducted analyses every year to determine the level of assessor reliability. These analyses indicate that assessors are making reliable, accurate, and fair evaluations of candidates' responses. The standards committees recommend to the National Board the specific standards for each certificate area and advise those involved in developing the corresponding assessment. The standards and the certificates are structured along two dimensions: the developmental level of students and the subject area.

Kansas Performance Teaching Portfolio

The Kansas Performance Teaching Portfolio (KPTP) requires teachers to provide information about the unit's lesson plans and assessments. Specific information about how the instruction is modified for two individual students within the classroom also is required. In addition, the teacher candidate reflects on the implementation of the unit for the whole class and the two focus students. The portfolios must address six focus areas:

- Analysis of contextual information
- Analysis of learning environment factors
- Instructional implementation
- Analysis of classroom learning environment
- Analysis of assessment procedures
- Reflection and self-evaluation

KPTP measures the teacher candidate's ability to design, deliver, and reflect on an entire unit of study through four distinct sources of evidence:

- Contextual information and learning environment factors
- Designing instruction
- Teaching and learning
- Reflection and professionalism

Training for Use of Product and Services

The available training for use of these products and services varies, depending on whether the state developed the rubrics in-house or used outside consulting services. The TQ Center's *Guide to Teacher Evaluation Products* provides additional information for each example.

Cost of Product and Services

The costs of these products and services vary, depending on whether the state developed the rubrics in-house or used outside consulting services. The TQ Center's *Guide to Teacher Evaluation Products* provides additional information for each example.

Advantages and Implementation Challenges

Advantages

- Evidence across various activities collected and considered by teachers, which encourages a perspective on teaching beyond the classroom.
- Potential for a more cohesive teaching team if the approach is applied collaboratively.
- May promote professional growth through provision of support to colleagues.

Implementation Challenges

- Time-sensitive feedback.
- Best when applied over the course of a year but difficult to regulate.
- Tension between using evidence as part of an evaluation and for professional growth.

How States Can Get More Information

- Washington ProTeach Portfolio:
- Alexandria (Virginia) City Public Schools Performance Evaluation Program (PEP):
- PACT Assessment—Teaching Event:
- National Board for Professional Teaching Standards:
- Kansas Performance Teaching Portfolio:

Research and Resources

Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved May 6, 2011, from

Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved May 6, 2011, from

National Comprehensive Center for Teacher Quality. (2010). *Guide to teacher evaluation products* [Website]. Washington, DC: Author. Retrieved May 6, 2011, from

APPENDIX E. TRIPOD SURVEYS: STUDENT, TEACHER, AND PARENT SURVEYS

Developer of Product and Services

The Tripod Surveys were developed by Ron Ferguson, Ph.D., at Harvard University, and Cambridge Education.

Description of Product and Services Available

Tripod surveys are one component of the Tripod Project, which aims to improve school capacity to address content, pedagogy, and relationships (the “tripod” of quality teaching) while closing achievement gaps. The surveys are available for students, teachers, and parents. Tripod surveys identify attitudes, perceptions, experiences, and practices in classrooms as they relate to the content knowledge of teachers, the pedagogical knowledge of teachers, and the relationships between teachers and students.

Tripod surveys examine the Seven C’s of quality teaching: care about students, control of student behavior, captivating students, clarifying lessons, challenging students academically, conferring with students, and consolidating knowledge. Tripod surveys are now in their 11th version. Previous research indicates that classrooms with high student ratings on the Seven C’s also produced higher average gains in student achievement. Currently, a modified version of the Tripod student survey is being used as part of the Measures of Effective Teaching (MET) Project funded by the Bill & Melinda Gates Foundation, which is researching the classroom practice of more than 3,000 teachers.

The Tripod student, teacher, and parent surveys were developed for use with teachers in any subject or grade level. The Tripod Project is now offering value-added analysis, using results from Tripod surveys to predict student achievement on state tests.

Training for Use of Product and Services

Resources and research on the Tripod Project can be found at the Materials Archive webpage

Cost of Product and Services

The Tripod Project offers consulting and support for student, teacher, and parent surveys; analysis and reporting; strategic school improvement planning; and professional development. Consultation services are customized based on client needs. For more information, see the Services and Offerings webpage

Advantages and Implementation Challenges

Advantages

- Can be used to report otherwise unobservable factors that may affect teaching, such as knowledge, intentions, expectations, and beliefs.
- Provides the unique perspective of the teacher.
- Provides the perspective of students, who have the greatest amount of experience with teachers.
- Can provide formative information to help teachers improve practice in a way that will connect with students.
- Makes use of the perspectives of students, who may be as capable as adult raters at providing accurate ratings.

Implementation Challenges

- Relies on teacher self-reporting, which may not be accurate.
- Should not be used as the sole or primary measure of teacher evaluation because student ratings have not been validated for use in summative assessment.
- Information on aspects of teaching (e.g., a teacher's content knowledge, curriculum fulfillment, or professional activities) not available from students.

How States Can Get More Information

States can find more information at the Tripod Project website
Rob Ramsdell (rob.ramsdell@camb-ed-us.com).

or by contacting

Research and Resources

Bill & Melinda Gates Foundation. (2010). *Student perceptions and the MET Project*. Seattle, WA: Author. Retrieved May 6, 2011, from

Ferguson, R. F. (2002a). Addressing racial disparities in high-achieving suburban schools. *NCREL Policy Issues*, 13. Retrieved May 6, 2011, from

Ferguson, R. F. (2002b). *What doesn't meet the eye: Understanding and addressing racial disparities in high-achieving suburban schools*. Cambridge, MA: Harvard University, Wiener Center for Social Policy. Retrieved May 6, 2011, from

ABOUT THE NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

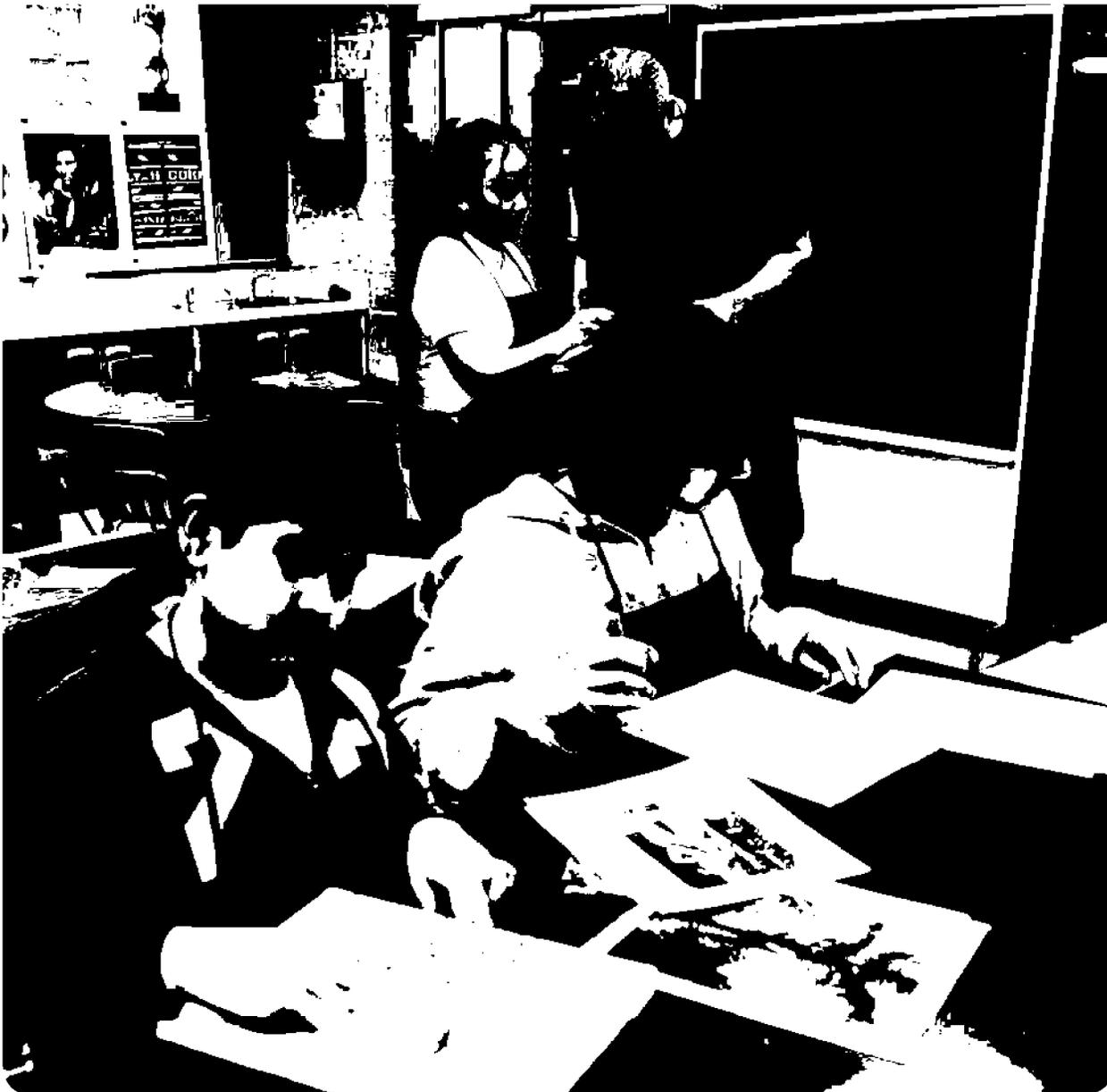
The National Comprehensive Center for Teacher Quality (TQ Center) was created to serve as the national resource to which the regional comprehensive centers, states, and other education stakeholders turn for strengthening the quality of teaching—especially in high-poverty, low-performing, and hard-to-staff schools—and for finding guidance in addressing specific needs, thereby ensuring that highly qualified teachers are serving students with special needs.

The TQ Center is funded by the U.S. Department of Education and is a collaborative effort of ETS, Learning Point Associates, and Vanderbilt University. Integral to the TQ Center's charge is the provision of timely and relevant resources to build the capacity of regional comprehensive centers and states to effectively implement state policy and practice by ensuring that all teachers meet the federal teacher requirements of the current provisions of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act.

The TQ Center is part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five content centers that provide expert assistance to benefit states and districts nationwide on key issues related to current provisions of ESEA.



Measuring Teachers' Contributions to Student Learning Growth for Nontested Grades and Subjects



ACKNOWLEDGMENTS

The TQ Center would like to thank the following individuals, who were instrumental in the development of this brief: Bill Slotnik, Community Training and Assistance Center; Margaret Heritage, Ph.D., National Center for Research on Evaluation, Standards, & Student Testing; Megan Dolan, Mid-Atlantic Comprehensive Center; Bonnie Billingsley, Ed.D., Virginia Tech; and E. Caroline Wylie, Ph.D., ETS.

Measuring Teachers' Contributions to Student Learning Growth for Nontested Grades and Subjects

This **Research & Policy Brief** was developed to help states consider options for assessing student learning growth for the majority of teachers who teach content not assessed through standardized tests.

March 2011

Laura Goe, Ph.D., *ETS*

Lynn Holdheide, *Vanderbilt University*



NATIONAL COMPREHENSIVE CENTER
FOR TEACHER QUALITY

CONTENTS

Introduction	1
Nontested Subjects and Grades	1
Measuring Growth	2
Why Measure Growth?	2
How Is Growth Measured?	3
Federal and State Priorities	4
Expectations for Teachers	5
Attribution and Student–Teacher Links	5
Factors for Consideration	7
Student Competencies in Specific Content Areas and Grade Levels	7
Identification of Reliable and Valid Assessments	7
Schoolwide Value-Added Models for Teachers of Nontested Subjects and Grades.	9
Measuring Student Learning Growth for Teachers in the Arts and Other Nontested Subjects.	13
Measuring Student Outcomes for “Caseload” Educators	13
Alignment With Federal Priorities	14
Application to All Grades and Student Populations	15
Standardized Evidence Collection	16
Measures That May Improve Teacher Performance	18
State Guidance to Districts	18
Comparability: Across or Within Districts?	18
Measures	19
Exceptions	19
Ongoing Research on Systems, Models, and Measures	21
Considerations for States: Moving Forward	21
Conclusion	22
References	23

INTRODUCTION

The growing need for more information about measuring teachers' contributions to student learning growth, particularly in nontested subjects and grades, is the impetus for this Research & Policy Brief. Although the research base in this area is disappointingly limited, the brief includes considerations and suggestions based on current models and experiences from the field. Although the brief is intended for use by states in developing statewide systems and providing guidance to districts, it also may be helpful to districts charged with designing and implementing evaluation models that fit within state and federal guidelines.*

For many states, the need to implement comprehensive teacher evaluation systems that consider teachers' contributions to student learning growth is clear and immediate. But because there are no research-based models for incorporating this component into teacher evaluation systems, states are experimenting with a variety of strategies to move forward. In fact, even without research to support particular approaches to evaluating teachers' contributions to student learning growth, states are proceeding—sometimes on very short timelines—to collect such evidence and incorporate it into a system of multiple measures of teacher performance. This endeavor is challenging even when there are standardized test scores that can be used as evidence of students' achievement progress, but it is especially complicated when no standardized measures exist, which is the case for the substantial percentage of teachers of nontested subjects and grades.

This Research & Policy Brief provides information about options for states to explore as well as factors to consider when identifying and implementing measures. The brief also focuses specifically on federal priorities to help ensure that evaluation systems meet the high

expectations set for teacher evaluation. Finally, the brief emphasizes the importance of fairly measuring *all* teachers, including them in the evaluation process, and ensuring validity in measurement.

Nontested Subjects and Grades

In *The Other 69 Percent: Fairly Rewarding the Performance of Teachers of Nontested Subjects and Grades* by Prince et al. (2009), “the other 69 percent” refers to the percentage of teachers whose contributions to student learning cannot be measured with test-based approaches (e.g., value-added models) because they teach subjects or grades that are not assessed with standardized tests.

Measuring effectiveness for the “other 69 percent” is probably the most challenging aspect of including student achievement growth as a component of teacher evaluation. According to Prince et al. (2009),

Identifying highly effective teachers of subjects, grades, and students who are not tested with standardized achievement tests—such as teachers of art, music, physical education, foreign languages, K–2, high school, English language learners, and students with disabilities—necessitates a different approach. It is important that states and districts provide viable options for measuring the progress of these groups of students and the productivity of their teachers, both of which contribute to school performance. (p. 1)

Statewide standardized testing is typically conducted for reading/language arts and mathematics in Grades 4–8 as required by the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act. Likewise, some states, albeit a smaller number, conduct such testing in certain grades for other subjects such as science

* See http://www.tqsource.org/webcasts/201012Workshop/Teacher_Effectiveness_Workshop_Glossary.pdf for a glossary of commonly used terms in current teacher evaluation reform efforts.

and social studies. Nontested subjects and grades in which standardized tests are not administered include the following:

- Subjects with standards that cannot be adequately or completely measured with a paper-and-pencil test (e.g., art, music, industrial arts, drama, dance)
- Subjects in lower elementary grades for which students cannot be reliably tested with paper-and-pencil or computerized tests (e.g., Grades K–2)
- Subjects/grades for which states have chosen not to test because of cost and priority relative to “core” academic subjects

In addition to nontested subjects and grades, there are certain student populations and/or situations for which standardized test scores are not available or utilized (e.g., students with cognitive disabilities). The Individuals with Disabilities Education Act of 2004 allows for the use of alternative assessments for students for whom the standardized assessment is inappropriate, even with reasonable accommodations. Moreover, smaller teacher caseloads for some student groups, such as students with disabilities and English learners, produce results that are statistically less reliable, often resulting in such groups being excluded in value-added or other growth models (Amrein-Beardsley, 2008; Feng & Sass, 2009).

Inclusion of teachers in nontested subjects and grades in an evaluation system that is based in part on teachers’ contributions to student learning growth requires the identification or development of appropriate measures and methods to accurately determine students’ growth toward grade-level and subject standards. Clearly, this task requires standards for every subject and/or grade level. If standards are nonexistent or poorly specified, it will be difficult to accurately determine teachers’ contributions toward growth in those subjects and grades, so ensuring that academic standards exist for every subject and grade should be a priority.

MEASURING GROWTH

Why Measure Growth?

Teachers are the most influential school-based factor on student achievement (Rivkin, Hanushek, & Kain, 2005; Sanders & Horn, 1998; Sanders & Rivers, 1996). Although studies have shown that some teachers are more effective than others at helping their students achieve at high levels, most indicators of teacher quality (e.g., credentials, characteristics, and observable practices) are generally poor predictors of student learning growth (Goe, 2007; Rice, 2003; Wayne & Youngs, 2003). Teachers’ scores on observation instruments have not been highly correlated with student learning growth (Weisberg, Sexton, Mulhern, & Keeling, 2009). However, it is not surprising that correlations are weak when the factors to be measured with observations are not well specified or when raters are poorly trained or inadequately monitored for scoring consistency after training.

Most of the indicators used in the past to determine teacher quality have been found to be inadequate, particularly when used in isolation, in differentiating between teachers whose students perform well and those whose students are not making adequate progress. Recent federal funding opportunities have emphasized teacher effectiveness and teacher evaluation based on teachers’ contributions to student achievement. This focus on evaluating teachers by measuring *student growth* rather than attainment is fairer to teachers whose students enter classrooms well below grade level. Teachers should not be penalized for choosing to teach in schools in which students are considerably behind their peers in proficiency. This is not to say that students’ mastery of appropriate grade-level standards is unimportant, but moving students as close as possible to proficiency, even if all students are not able to reach it, should be the focus of teachers’ efforts. Teachers should be given

credit when these efforts succeed, and using multiple measures of student learning growth is essential to ensure that teachers in all subjects and grades are fairly credited.

How Is Growth Measured?

Since the initial passage of ESEA, standardized assessments have been used to determine student progress toward academic standards. Value-added models and other growth models have generated considerable interest for showing growth over time for students, and lately, for the teachers of those students. Recent efforts to create statewide longitudinal data systems that link teachers with their students' achievement have set the stage for states and districts to use student learning growth on standardized tests as part of determining teacher effectiveness. However, in most states, only reading/language arts and mathematics in Grades 4–8 are actually tested with state standardized assessments, meaning that teachers in most subjects and grades do not have state standardized test results that can be used as components of teacher evaluation.

How results from standardized tests are actually used as part of teacher evaluation remains an open question because states and districts are just beginning to use linked student–teacher data and growth models, (e.g., value-added models). Tennessee is at the forefront of these efforts because it has been using the Tennessee Value-Added Assessment System (TVAAS) for more than a decade to provide individual teachers and their principals with the teachers' district rank based on value-added measures. Many more states are developing systems that will allow them to use growth models such as EVAAS (the version of TVAAS that is not state-specific) as well as the Colorado Growth Model, which focuses on students' growth toward proficiency (See "Different Approaches to Measuring Students' Growth"; Betebenner, 2008).



DIFFERENT APPROACHES TO MEASURING STUDENTS' GROWTH

Although most teachers currently cannot be evaluated with growth models based on standardized tests, it may be helpful to understand how growth models might fit within an evaluation system. A number of states are planning to implement (or already have implemented) value-added or other types of growth models. In its simplest form, the value-added measure as it is used for evaluating teachers is calculated as follows: Students' previous test scores are used to create predicted test scores for a given year. The difference between the predicted and actual test scores are growth scores. Teachers' contributions to student learning are determined by calculating the average of all of their students' growth scores. The teachers are then ranked with other teachers within a district (or other unit of interest) according to how much they contributed to student growth, and this ranking is their value-added "score."

In some value-added models, only students' prior achievement scores are used in the calculation; other models include students' gender, race, and socioeconomic background; still others include information about teachers' experience. With a value-added measure, teachers whose students performed as well as predicted are considered "average" teachers; those whose students performed better than predicted are considered "above average" or "highly effective"; and those whose students performed worse than expected are considered "below average."

The Colorado Growth Model focuses instead on student growth percentiles. Students are compared with their academic peers (i.e., students at the same starting point in achievement) to determine normative growth. The goal is to determine students' standing relative to their academic peers. Thus, if students' scores are better than those of their academic peers, they are performing well. All of a teacher's students can be scored in this way, resulting in an average growth for the class or the teacher's roster, which can then be attributed to the teacher's efforts in much the same way value-added scores are.

Whenever such models—whether value-added models, the Colorado Growth Model, or other models—are used, results should never be considered in isolation as the sole measure of a teacher's performance but rather included in a system of multiple measures that produces a comprehensive picture of a teacher's performance.

However, results obtained through such growth models have rarely—until now—been used as part of teacher evaluation. Even in those states that have the capacity to collect such information, questions remain about the accuracy of the information, given evidence of year-to-year fluctuation in teachers' scores (Braun, Chudowsky, & Koenig, 2010; Koedel & Betts, 2009; McCaffrey, Sass, Lockwood, & Mihaly, 2009; Schochet & Chiang, 2010).

For teachers in nontested subjects and grades, there are few state models that demonstrate how contributions to student learning growth can be systematically measured and analyzed in ways that allow for differentiation among teachers. Some experiments are currently under way in collecting evidence of student learning growth for these teachers, but research has not yet been conducted on how such evidence is being used within evaluation systems.

Federal and State Priorities

To position themselves for a successful Race to the Top bid, many states passed new legislation mandating that student achievement growth be included as part of teacher evaluation. Federal priorities (Secretary's Priorities for Discretionary Grant Programs, 2010) specify that acceptable measures for determining teachers' contributions to student learning must meet the following requirements:

- Rigorous
- Between two points in time
- Comparable across classrooms

These terms are not explicitly defined in Race to the Top guidance. In fact, the federal government has declined to offer definitions for these terms, preferring instead to encourage states to define them locally. For federal purposes, Race to the Top winners must follow

through with what they promised in their plans, which may include defining terms. The following considerations may provide some assistance in the development of state definitions:

- **Rigorous** measures may exhibit high expectations for student progress toward college- and career-readiness. In other words, an assessment that measures student progress in social studies would be designed to measure students' mastery of grade-level standards for that subject. Thus, a student who does well on such an assessment should be on track to successful, on-time promotion to the next grade and ultimately to graduation.
- **Between two points in time** may mean assessments that occur as close as possible to the beginning and end of a course so that the maximum growth toward subject/grade standards can be shown.
 - *Example:* An Advanced Placement (AP) test may serve as an end point, but another assessment (aligned with the state standards and focused on the specific knowledge and skills measured by the AP tests) will likely need to be administered at the beginning of the year to establish students' level of mastery of the standards when they begin the course to determine teachers' contributions to student growth. The process of collecting evidence of students' initial skills and knowledge should not be undertaken lightly. Ideally, an assessment that has been designed and created by experts specifically to serve as a pretest should be used.
 - *Example:* Student portfolios representing mastery of standards could be collected at the end of the year. However, at the beginning of the year, teachers would need to collect and score evidence (i.e., activities or assessments aligned with the state standards and focused on

the specific knowledge and skills needed for creating a successful portfolio) that would allow them to formulate an initial score point for each student. Through this process, increased knowledge and skills could be documented for individual students.

- **Comparable across classrooms** has two possible interpretations, both of which are useful to consider:
 - The measures used to show students' growth for a particular subject are the same or very similar across classrooms within a district or state.
 - The measures used in *nontested* subjects and grades are as rigorous as those in tested subjects and grades. In other words, measures used to document student learning growth in art, music, and social studies must be as rigorous as those for student learning growth in reading/language arts and mathematics.

Expectations for Teachers

Race to the Top defined an *effective teacher* as one whose students achieved at least one grade level of academic growth during the course of the year and a *highly effective teacher* as a teacher whose students achieved at least one and a half grade levels of academic growth during that time frame. Although not federally mandated, teachers are generally required to ensure that their students are on track to meet grade-level expectations. In addition, they are expected to regularly evaluate student progress and issue grades that reflect students' efforts and achievement in mastering the content. With new federal and state mandates calling for the inclusion of teachers' contributions to student learning in the evaluation process, growth must be documented in some way, which means that teachers in nontested subjects and grades need to focus on

new approaches to measuring their students' progress—approaches that are rigorous, that provide data on growth between two points in time, and that are comparable across classrooms.

Attribution and Student–Teacher Links

Determining teacher attribution for particular students is challenging. What if a student receives services in a general education classroom in which coteaching occurs? Should both teachers be held accountable for student growth? How will paraprofessionals' contributions to student learning growth be sorted out from those of the content area or special education teachers?

In a recent TQ Center inquiry, 85 percent of the local and state special education administrators polled were of the opinion that both the general and special education teachers should be held accountable for all students in the class (Holdheide, Goe, Croft, & Reschly, 2010). However, there may not be widespread agreement for that approach. Linking student growth (or a portion thereof) to the appropriate teachers presents challenges.

One approach developed by the Ohio-based Battelle for Kids is the use of new linkage software that has the capacity to account for student mobility and shared instruction/coteaching in subject areas for which value-added data are available (See “Student–Teacher Linkage for Attribution”). This approach also may be viable using other types of student growth measures, as it facilitates a deeper and often necessary discussion regarding teacher roles and responsibilities. At this time, however, a research-based methodology for this type of teacher-led determination has yet to be established. In addition, its application in a non-value-added growth measure needs to be explored.

Teacher apprehension toward accountability systems including student growth measures can be minimized if teachers perceive the system to be fair and accurate. For example, failure to directly address which teachers are accountable for which students will likely result in pushback from teachers. In addition, teachers need to have an opportunity to verify their rosters of students and the length of time that students were on their rolls. This verification process is particularly important in schools with high rates of absenteeism or student mobility. Teacher involvement and support in this process is essential. Teachers must be involved in the processes of problem-solving, collecting data during implementation, and obtaining feedback on effectiveness. Teachers know their classrooms, their students, and the way in which they collaborate with other teachers.



STUDENT-TEACHER LINKAGE FOR ATTRIBUTION

Olentangy Local School District in Ohio and other districts across the country are taking value-added analysis to the classroom level with Battelle for Kids' innovative, Web-based BFK·Link™ solution to accurately “link” teachers to students. During the linkage process, teachers review and correct data used for teacher-level measures of effectiveness, including value-added analysis, by ensuring that all students taught are “claimed” by teachers for all subjects, accounting for student mobility and shared instruction/coteaching.

The BFK·Link process attempts to maximize correct matching of teacher effort to student outcomes through a transparent process. For example, for teachers working in a true coteaching situation, both teachers may each “claim” 50 percent of each student. Or, if students receive some support services in a resource room, the general educator may claim 70 percent while the special education teacher claims 30 percent. Student standardized test scores are then linked with teachers for the percentages specified.

In typical classrooms, teachers claim 100 percent of most of their students, with reduced percentages for students with special needs who receive services from other teachers. The system verifies accuracy by marking cases in which a student has more or less than 100 percent for inspection (i.e., more than one teacher is contributing to that student's scores, but the teachers' combined percentages do not add up to 100), and the teachers are asked to reevaluate. When percentages add up to 100 percent, the BFK·Link solution calculates scores proportionally.

The use of value-added analysis to inform instruction and high-stakes decisions requires accurate linkage of teachers to students. For more information, see *The Importance of Accurately Linking Instruction to Students to Determine Teacher Effectiveness* (Battelle for Kids, 2009), a white paper commissioned by the Bill & Melinda Gates Foundation.

FACTORS FOR CONSIDERATION

States and districts attempting to incorporate student growth into their teacher evaluation systems are faced with the challenge of identifying other valid and reliable measures for teachers of nontested subjects and grades. Though the research base is still developing, the following questions may be useful to consider during the problem-solving process:

- Is there a consensus on the competencies students should achieve in this content area?
- What assessments/measurements can be used to reliably measure these competencies with validity?
- Should the use of schoolwide value-added models be considered as a means to measure student progress in nontested subjects and grades?
- How will growth in performance subjects (e.g., music, art, physical education) be determined?
- How will related personnel (“caseload” educators) be factored into the system?
- Do these measurements meet all of the federal requirements (i.e., rigorous, between two points in time, and comparable across classrooms)? Are measurements aligned with federal priorities?
- Can these measurements be applied to all grades and student populations?

Student Competencies in Specific Content Areas and Grade Levels

In most states, content standards are designed by a group of experts and practitioners to encourage proficiency for every student by defining the knowledge, concepts, and skills students should acquire for each subject. Each standard typically has clearly defined

statements and examples of what all students should know and be able to do at the end of a particular grade. These standards often drive changes in certification, assessment, curriculum, instructional strategies, and teacher professional development. Therefore, a transparent alignment to these content standards offers guidance when identifying and/or designing assessments to measure student progress, which could be used to determine teachers’ contributions for evaluation purposes. In states in which subject content standards exist, these standards provide a basis for the identification and development of assessments.

Identification of Reliable and Valid Assessments

States are struggling most with determining appropriate measures for evaluating teachers’ contributions to student learning growth in the nontested subjects and grades. The challenge facing many states, including the Race to the Top award recipients, is to identify valid, reliable processes, tools, assessments, and measures that allow them to collect data to measure every teacher on his or her contributions to student learning growth. Many current approaches to measuring teachers’ contributions to student learning in the nontested subjects and grades do not meet all of the federal criteria of rigor, comparability, and growth measured across two points in time.

Local and state education systems have taken various approaches, each of which has its own strengths and limitations as indicated in Table 1. None of these options is “perfect,” and concerns about validity, reliability, and costs are associated with nearly all of them. The trade-offs involved with using these measures should be considered by stakeholder groups as well as state and district evaluation and assessment personnel.

Table 1. Options for Measuring Student Growth to Inform Teacher Evaluation in Nontested Subjects and Grades

<p>Use existing tests designed for other purposes, such as end-of-course tests that may be included with some curriculum packages.</p>	<p>Tests developed by the creators of the curriculum are likely to be aligned well with the content of the course.</p> <p>It may be possible for the creators of the curriculum to develop appropriate pretests if they are not included in the package.</p>	<p>Validity is a concern whenever a measure is used in a way that was not intended by the maker of the assessment (e.g., turning end-of-course assessments into pretests). Discussions with the test maker about using tests for other purposes may provide insight into how validity may be affected.</p>
<p>Create new tests for areas in which few assessments exist.</p>	<p>Tests can be developed in alignment with specific grade/subject standards.</p>	<p>This option is a costly undertaking, given how much effort goes into developing valid and reliable tests that can accurately measure students' knowledge and skills based on a set of subject/grade standards.</p> <p>Paper-and-pencil tests may not be appropriate as the sole measure of student growth, particularly in subjects requiring students to demonstrate knowledge and skills (e.g., art, music).</p>
<p>Use the four Ps—portfolios, products, performances, or projects—to measure student growth over time for subjects in which standards require students' to demonstrate mastery.</p>	<p>Evidence about student growth in particular knowledge and skills can be documented over time using performance rubrics.</p> <p>Portfolios and projects reflect skills and knowledge that are not readily measured by paper-and-pencil tests.</p>	<p>Training would be required for everyone involved in using rubrics to ensure reliability (i.e., all raters agree on how the evidence reflects different levels of achievement).</p> <p>Performance ratings are best conducted by groups of raters rather than individual teachers; bringing raters together to examine student work may be a logistical challenge.</p>
<p>Give teachers in nontested subjects and grades a “prorated” score for collaboration with a teacher in a tested subject (i.e., an art teacher collaborating with a mathematics teacher).</p>	<p>No additional resources are required. This option is similar to the Teacher Advancement Program (TAP) model.</p>	<p>Determining prorated scores would be problematic, threatening the validity of the information.</p> <p>Differences among methods of determining contributions of these collaborating teachers may make it difficult to ensure comparability.</p>
<p>Use other measures (e.g., classroom observations) for these teachers.</p>	<p>No additional resources are required.</p>	<p>No information about student achievement is obtained, meaning that this option will not meet federal priorities and many state requirements.</p> <p>Observations and other measures focused on teacher practice offer little information about students' actual achievement in a teacher's classroom.</p>
<p>Use student learning objectives (i.e., the teacher selects objectives and determines how to assess student growth toward meeting objectives).</p>	<p>Teachers benefit from being directly involved in assessing students' knowledge and skills.</p> <p>Teachers can set learning objectives based on students' special needs (e.g., students with disabilities or English learners).</p> <p>This option is applicable to all teachers and subjects.</p>	<p>Comparability across classrooms will be problematic because of teachers' selection of assessments and objectives.</p> <p>This option is very resource-intensive for principals or district personnel who approve objectives, provide teachers with guidance, verify outcomes, and so on.</p>

Schoolwide Value-Added Models for Teachers of Nontested Subjects and Grades

The use of schoolwide value-added scores has been suggested as a way to evaluate teachers in nontested subjects and grades to remedy the lack of available measures. Similar to the Teacher Advancement Program (TAP) model, it is perhaps the least expensive method of including these teachers in a test-based evaluation system because new measures and teacher training are not required. In this scenario, teachers of nontested subjects would be given the schoolwide value-added average in place of individual growth results.

This approach presents some additional challenges for a number of reasons, including questions about rigor and comparability when judgments are made about individual teacher performance based on students they never taught. Furthermore, it is much more difficult to learn about teachers' contributions to student achievement if they are assigned scores based on other teachers' efforts. Mathematics and reading/language arts value-added information will not be useful to teachers in improving their performance in subjects such as art, social studies, and science. In addition, failing to measure progress in these subjects and for certain students devalues the contributions those teachers make to student learning and provides no information about their effectiveness in teaching their subject matter.

Using Existing Assessments

In the search for measures to determine teachers' contributions to student learning growth, it is likely that an iterative process will be needed. After a potential instrument is identified, it is necessary to demonstrate that the measure is valid for the intended purpose (i.e., that the measure does, in fact, differentiate among teachers whose students have high levels of learning growth and teachers whose students' learning did not increase

at acceptable levels). Because the measures that might be used for teacher evaluation have not been validated for this purpose, it is important to analyze data collected by using these measures and determine whether the data show differences among teachers and whether results from using these measures correlate with other measures in the evaluation system.

The validation process generally starts with determining the factors that need to be measured and for what purpose. As part of this process, it is important to consider the evidence needed to measure teachers' contributions to student learning growth. Evidence will have been gathered to build a case for using a particular measure as part of the evaluation system (Herman, Heritage, & Goldschmidt, in press). After the types of necessary evidence are determined, measures and instruments that can be used to collect such evidence must be identified. Then, results from using measures must be analyzed to determine how the measures performed in practice.

For example, if the district wanted all Grade 8 reading/language arts teachers to administer an essay to students at the beginning and end of the year to establish student growth, the district would need to score (or preferably have teachers score together) the essays and determine whether they show student learning growth. A distribution of scores would need to be made and cross-referenced with teachers to determine whether more or less growth occurred in particular teachers' classrooms or the pattern of growth is random. A random pattern would suggest that the growth students made was not necessarily attributable to a particular teacher's efforts, whereas a pattern of higher or lower growth associated with a particular teacher may be an indicator of his or her efforts. Comparing these results with results from additional measures (e.g., other assessments, projects, portfolios) should then be helpful in validating the usefulness of the essays in showing teachers' contributions to student growth.

In addition, validity is a matter of degree—it is seldom perfect, but a high degree of validity must be achieved when results will be used for high-stakes purposes such as teacher tenure, performance pay, and dismissal. Clearly, the higher the stakes, the greater validity is needed in terms of the evidence. In addition, validity can be improved over time by identifying which measures are and are not working to provide evidence to make decisions about teacher performance.

For most states and districts, waiting until the measures are perfected may be impractical, given the timelines to implement new teacher evaluation systems. So even though the measures may have weak evidence of validity in the first attempts at implementation, states and districts will benefit from creating a process to continually evaluate and strengthen the measures or eliminate those that continue to show weak evidence of validity. Over time, a collection of measures with strong evidence of validity will be created. Obviously, this process is neither quick nor easy, and it requires some expertise. Districts and states with limited capacity may consider joining forces with others in the region to share resources rather than “reinventing the wheel” in each district or state.

Utilizing existing assessments and avoiding the development of new assessments certainly holds appeal for implementation ease. Interim or benchmark assessments are already widely used in schools as a means to provide assessment of student progress toward content standards. In fact, schools that implement response to intervention (RTI) have likely identified measures for the progress monitoring component of implementation. These assessments are often embedded into the instructional cycle and are used to make the necessary instructional adjustments to facilitate student mastery. Working collaboratively

with state and district RTI initiatives to identify potential sources of evidence for evaluation purposes may facilitate a combined effort to address the persistent achievement gaps in schools (See “National Center on Response to Intervention Progress Monitoring Tools Chart”).



The National Center on Response to Intervention annually publishes a progress monitoring tools chart to assist educators in identifying tools that best meet their needs. The Center’s Technical Review Committee (TRC) independently established a set of criteria for evaluating the scientific rigor of progress monitoring tools.

Included in this chart are ratings for instrument reliability of the performance-level score, reliability of the slope, validity of the performance-level score, predictive validity of the slope of improvement, and disaggregated reliability and validity data. In addition, the charts include the standards by which the TRC reviewed each tool (e.g., whether the tool is available in alternative forms, its sensitivity to student improvement, and its ability to measure end-of-year benchmarks).

This chart can be accessed at http://www.rti4success.org/tools_charts/progress.php.

Although these existing assessments were not designed specifically to inform teacher evaluation, they may have merit for that purpose. However, it is not as simple as adopting existing assessments. A thorough review of each assessment should be conducted, including its validity in measuring progress on the specific content standards and its measurement reliability across students and teachers. Moreover, assurance that these assessments measure what is valued is essential if evaluation results will be used to make personnel and compensation decisions.

Examples of Approaches to Assessment

Hillsborough County, Florida. Hillsborough County, Florida, a recent Race to the Top award recipient, has taken the approach of developing new assessments specifically designed to assess content mastery and plans to use data to inform teacher evaluation. Each nontested subject will have a pretest and posttest in which student scores are averaged over a three-year period to determine teacher effectiveness. As indicated in Table 1, this approach is fairly time and cost intensive; however, newly developed end-of-the-course assessments are more likely to be readily aligned with the content standards and have the potential to meet two of the federal requirements: *comparability* and *across two points in time*. Compliance with rigor would be dependent on how the data are used to determine acceptable student growth, and therefore, teacher proficiency.

Delaware. The state of Delaware uses a combination of approaches in which existing and new measurements are identified, assessed, and determined to be acceptable by experts at the state level. With the assistance of trained facilitators, Delaware assembled a group of local practitioners, arranged by content area expertise, to conduct a thorough review of existing measurements. After consensus was reached, the group submitted to the state a listing of recommended assessments and/or methods to assess student growth toward the content standards. This listing is updated and shared regularly (after approval from an independent panel of experts).

Austin, Texas. States also may identify specific criteria required for assessments to be considered valid measures of student growth. In Austin, Texas, teachers participating in a pay-for-performance pilot are involved in determining student achievement growth through the development of student learning objectives (SLOs). SLOs are classroom, grouping, or skill-based objectives, and teachers' ability to meet the SLOs determines their level of effectiveness. The quality of SLOs in measuring student growth is established by a rubric that determines whether the objectives and associated assessments are rigorous, measureable, reliable, and valid and whether the projected growth trajectory is considered rigorous. Although this approach facilitates teacher investment in the process, which is a definite strength, maintaining rigor is dependent on the rubric's implementation fidelity among administrators and teachers. In addition, SLO results may not be comparable across classrooms because various assessments are used to establish student growth. Moreover, if the evaluation system includes observations conducted by administrators, the burden on the administrators may be substantial.

For more information about these assessment approaches, see "Practical Examples of State Evaluation Systems."

PRACTICAL EXAMPLES OF STATE EVALUATION SYSTEMS

Hillsborough County Public Schools, Florida

Hillsborough County is the recipient of a seven-year, \$100 million Bill & Melinda Gates Foundation grant and has recently been awarded Race to the Top dollars to continue its efforts to improve results through the Empower Effective Teachers (EET) program.

The goals of EET are to:

- Develop a quality induction program for new teachers.
- Improve the teacher and principal evaluation system.
- Enhance the system of professional development.
- Provide effective incentives for teachers and improve the compensation plan.

Hillsborough County uses multiple measures to determine teacher effectiveness including peer and principal ratings using a modified version of Charlotte Danielson's *Framework for Teaching*. Those ratings make up 60 percent of teacher evaluations, with student performance on the Florida Comprehensive Assessment Test or end-of-course examinations making up the remainder.

Hillsborough County's stated commitment is to evaluate every teacher's effectiveness with student achievement growth, even teachers in nontested subjects and grades. To do so, Hillsborough County is in the process of creating pretests and posttests for all subjects and grades, expanding state standardized tests, and using value-added measures to evaluate more teachers.

In the 2010-11 school year, the statewide assessment program began transitioning to assessing student understanding of the Next Generation Sunshine State Standards through the implementation of the Florida Comprehensive Assessment Test® 2.0 (FCAT 2.0) and Florida End-of-Course Assessments.

Information on Hillsborough County's EET program can be accessed at <http://communication.sdhc.k12.fl.us/empoweringteachers/?p=611>.

Delaware

Delaware already had an excellent statewide evaluation system, which required classroom observations and encouraged teachers to focus on school, district, and state goals as well as their own professional growth. Delaware conducted a yearly external evaluation of the system, soliciting feedback from teachers and administrators through surveys, interviews, and focus groups. Revisions were made to the system yearly based on these results. The state also collaborated with the teachers union to ensure that evaluations were fair and responsive to the needs of the teachers and administrators. However, Delaware's system was lacking a mechanism to evaluate teacher contributions to student learning growth.

One reason that the state was awarded Race to the Top funds was the collaborative nature of the proposal, bringing stakeholders to the table at every step. As state staff focused on implementation, they continued to involve stakeholders in each step of the discussions. They valued teacher and administrator input, which was reflected in the steps they took to identify appropriate measures for the nontested subjects and grades as well as additional measures for teachers whose students took the state standardized test. A team of trained facilitators led groups of teachers as they met to discuss measures they currently used to evaluate their students' growth toward grade/subject standards. After discussing the merits of the measures and how they could be used, teachers made recommendations to the state about which measures to include.

The TQ Center and the Mid-Atlantic Comprehensive Center have been partners with Delaware during the implementation of its Race to the Top plans. In addition, Delaware has sought assistance from the Assessment and Accountability Comprehensive Center in convening a panel of experts to evaluate the potential measures for statewide use to show teachers' contributions to student growth in various grades and subjects. This process is ongoing.

Austin Independent School District Reach Compensation and Retention System, Texas

The Austin Independent School District Reach Compensation and Retention System is a four-year pilot incentive pay program for teachers and principals initiated in 2007–08. The program goals are to:

- Ensure quality teachers in every classroom.
- Provide professional growth opportunities.
- Increase retention.

The program focuses on student growth, professional growth, and schools with the highest need. Student growth is measured by student learning objectives (SLOs). Each teacher develops two SLOs—one that targets classroom performance and the other focused on a particular skill or subgroup of students (e.g., students with special needs). Each SLO must be a measurable objective that is approved by the principal. Teachers and principals undergo a series of trainings on how to establish and measure learning objectives.* The SLO's appropriateness, rigor, and acceptability are determined through the use of a rubric that considers the following questions:

- What are the needs?
- What and who is targeted?
- What will students' learn?
- How will you know whether they learned it?
- What is your goal for student achievement?
- How rigorous is your SLO?

Information regarding this system and the rubric can be accessed at <http://www.austin.isd.tenet.edu/Inside/Initiatives/compensation/releases.phtml>.

*SLOs are used to determine incentives and are not an integral part of the evaluation of teachers at this time.

Measuring Student Learning Growth for Teachers in the Arts and Other Nontested Subjects

Not all standards can be adequately assessed with a multiple-choice test. Many subjects require students to perform or create a product to demonstrate mastery of the standards. For these subjects, one or several of the four Ps (i.e., portfolios, performances, products, and projects) will likely be required to assess music students' ability to play scales on their chosen instruments; art students' ability to create works of art in various mediums; foreign language students' ability to speak the language they are studying; and family and consumer science students' ability to budget, plan, and prepare a wholesome family meal.

For these subjects, the focus is on designing appropriate tasks (e.g., performance, activities) that demonstrate students' mastery of standards and then developing appropriate pretests that allow districts/schools to determine students' knowledge and skills at the *beginning* of the course. In some cases, students can perform the *same* task: music students' can play the same piece of music at different points in time to show progress; art students can draw a still life; drama students can perform a monologue; and so on. In other cases, it may not be feasible for students to perform the same task. In these instances, it may be useful to identify the specific knowledge and skills that students need to know to successfully demonstrate mastery of a particular standard and then identify or develop tasks to serve as pretests from which progress on those standards can be determined.

Measuring Student Outcomes for “Caseload” Educators

Not every educator has a classroom. And some educators are responsible for services delivered to the entire school, not just a class. These related personnel (e.g., counselors, school psychologists, librarians, school

nurses, and speech therapists) may work with individuals but also with small or large groups of students. Although many states do not require the evaluation of such personnel in parallel with teachers, these “caseload” educators are included in the educator evaluation system in a number of states and districts. To measure their contributions to student learning growth, it may be helpful to think of them as having “caseloads.” For example, a school counselor may have a caseload that includes:

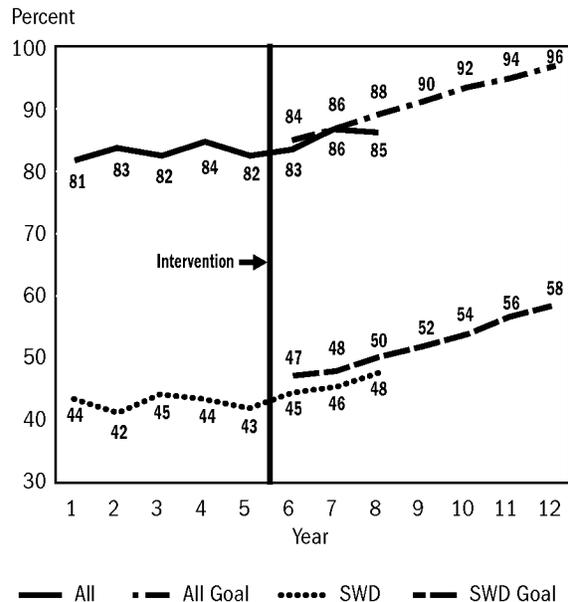
- All the students in the school (i.e., providing services such as career counseling at the high school level).
- Students experiencing emotional or behavioral problems.
- Students in crisis because of family events or relationship issues.
- Students with frequent unexcused absences.
- Teachers (e.g., providing professional development on recognizing the signs of physical or sexual abuse and what the law requires them to do).

Caseload educators may not be directly involved with academic content, making determining their contribution to academic achievement more difficult. These personnel may want to document their contributions to growth in terms of both educational successes and other types of outcomes. For example, a high school guidance counselor may want to track the proportion of students enrolling in AP classes, the proportion of students engaging in extracurricular activities, or the proportion of students for whom attendance rates have increased.

Caseload educators, and their associated goals, will likely vary according to the discipline and needs at the school, building, classroom, group, or individual student level. For example, a school with attendance issues may concentrate on attendance, whereas others may turn their attention toward AP course enrollment, reduction in incidences of bullying, or increased interactions between educators and parents.

Documented progress toward goals can be charted and monitored on an Excel spreadsheet, as illustrated in Figure 1. Likewise, intervention implementation can be tracked and monitored to determine effectiveness.

Figure 1. Sample of Documented Progress for Student Attendance



Source: Reschly and Holdheide (2010)

Alignment With Federal Priorities

Some measures are more likely than others to comply with federal priorities and state legislative mandates; however, these various approaches generally lack supporting research, leaving states and districts to their own devices to determine which options are most feasible. State and district priorities, financial resources, human capacity strengths and limitations, professional development needs, and system capacity issues should be contemplated prior to making decisions. General guidelines for selecting measures include the following:

- Avoid “reinventing the wheel.” If tests already exist that can be used for measuring teachers’ contributions to student learning, consider them first and determine whether they are useful in differentiating among levels of teacher effectiveness.

- Evaluate the available evidence for using the assessment as a measure of student growth for teacher evaluation.
 - Continue to evaluate the evidence by collecting and analyzing data resulting from the use of particular measures, including correlating measures with each other.
- Focus on measures that meet federal and state requirements and priorities by putting them to the following test:
 - Measures *must* show students' growth "between two or more points in time."
 - Measures *must* be "comparable across classrooms."
 - ◆ Consistency of measures across all teachers in a grade/subject ensures comparability of results.
 - ◆ For the four Ps—portfolios, products, performance, and projects—common rubrics should be used and agreement should be established as to how they will be used and who will score them.
 - Measures *must* be "rigorous."
 - ◆ Measures must be based on appropriate grade-level and subject standards.
 - ◆ Measures must demonstrate high expectations for student learning (i.e., on track to produce college- and career-ready graduates).
- Involve teachers and administrators in decision-making processes. They will benefit from their involvement, and their participation in considering appropriate measures will ensure greater "buy-in" for the results of the process.
- Choose measures that have the potential to help teachers improve their performance by:
 - Motivating teachers to examine their own practice against specific standards.
 - Allowing teachers to participate in or co-construct the evaluation (e.g., "evidence binders").
 - Giving teachers opportunities to discuss the results with evaluators, administrators, colleagues, teacher learning communities, mentors, and coaches.
- Choose measures that are directly and explicitly aligned with:
 - Teaching standards.
 - Professional development offerings.
- Include protocols and processes that teachers can examine and comprehend.

Application to All Grades and Student Populations

Assessing the effectiveness of teachers of students with disabilities and English learners presents challenges to determining teacher effectiveness due to the unique and varied roles these teachers assume (Holdheide et al., 2010). Likewise, measuring growth using standard measures for students with disabilities can be problematic, as standards-based models to determine growth are not based on individualized student goals.

The general tendency is to identify a different system or set of measures for special education teachers or English language specialists. Students with special needs and English learners have varying levels of ability and are taught in many different settings (e.g., general education classroom, resource room, separate classroom). Therefore, the types of assessment used to determine student growth may vary depending on the curriculum taught in the specified setting. Many students with special needs receive services in the general education classroom in which the assessments for determining student growth could (or should) be the same (possibly with accommodations) as that of students without disabilities, especially if these measures are vertically equated. For example, states may use the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) assessment (Good & Kaminski, 2002, 2011) to determine student progress in reading

and the effectiveness of teachers in teaching reading, particularly if the state does not have a standardized measure of reading in early grades. The DIBELS assessment would be appropriate for general education students, including students with disabilities who are participating in the general education curriculum.

The appropriateness of each content-specific or grade-specific assessment should be considered, and appropriate accommodations should be provided as needed. Similarly, some students with disabilities are working toward alternative standards, such as a life skills curriculum, which is not reflected in the standardized tests. In this scenario, different assessments need to be identified in order to measure student growth toward those alternative standards. Therefore, participation by teachers of students with disabilities is essential as states assemble teams to design and develop appropriate measures in all achievement areas included in the standard curriculum. Special education teachers who serve in inclusion models and engage in coteaching are able to bring a perspective to this work that addresses the needs of general and special education students, thereby contributing to the design of appropriate assessments in the areas not currently tested with standardized measures. Separate teams of special educators who instruct toward alternative standards also may be developed, as their measures would vary considerably due to content and ability level.

Student progress on the individualized education program (IEP) has emerged as a potential source for measuring teacher effectiveness for students with disabilities. In one sense, it is not surprising because most IEPs contain individualized goals that are aligned with state standards, including measureable objectives that are monitored regularly for student progress. However, IEPs were never intended to be used as a tool to measure teacher effectiveness, and using them this way likely will raise legal and other potentially contentious issues. Though the

individualized nature of the IEP and the detailed description of present levels and objectives for growth are positive features, standardized measures based on the general curriculum are still needed to assess teacher effectiveness.

STANDARDIZED EVIDENCE COLLECTION

Many states and districts are attempting to build comprehensive teacher evaluation systems that are responsive to federal priorities but are finding that there is little research to support the use of particular systems, weights, or measures. Because few states and districts currently have evaluation systems that incorporate multiple measures, there has been little opportunity to conduct research on how these measures perform. The question remains: Do the various measures in some weighted combination accurately identify teachers at different levels of effectiveness? Until systems with multiple measures and various weighting schemes are employed over time and evaluated by researchers, states and districts must be guided by general knowledge about how to use measures in a way that yields results that are rigorous and comparable.

One general method to ensure greater rigor in how multiple measures of all types are used is to implement *standardized* evidence collection. Everyone is familiar with the term *standardized test*. A standardized test is a test that is given according to specific rules that ensure that the test results will be comparable across students, schools, and districts. Specific rules also can be created and followed for all types of measures. By standardizing evidence collection, greater comparability across teachers is possible. Table 2 offers some suggestions for standardizing evidence collection for different types of measures of student learning growth.

Table 2. Standardizing Evidence Collection for Different Types of Measures

Curriculum-based pretests and posttests	Ensure that all teachers give the tests on the same day at the same time and allow students the same amount of time for completion. Teachers should agree to limitations on test preparation for posttests.	Accurately determining growth may be difficult in schools where students are particularly advanced versus schools where students begin the year below grade level. Adjustments may need to be made to account for these differences. Some students may do very well on the initial pretest, making it impossible to show growth. Providing those students with additional challenging curriculum and enrichment activities may allow them to show growth.
Student portfolios	Engage all teachers who plan to use student portfolios in the process of determining what constitutes acceptable evidence for various levels of performance (i.e., characteristics of a “beginning” versus “advanced” still life drawing). Develop or adopt appropriate rubrics and forms for teachers to use in establishing students’ beginning performance levels on the knowledge and skills needed to meet the grade/content standards reflected in the portfolio. The same rubrics and forms can be used to evaluate the portfolio at the end of the course.	Portfolios should include not only the students’ work but also the teachers’ scoring rubric and comments and the students’ reflections (i.e., how the student plans to improve upon the work). They should not be a catch-all for multiple iterations of an essay or other unrelated work. Teachers need to work together to create or adopt a rubric and scoring approach to ensure that they all agree on the characteristics of a “beginning” versus “advanced” effort. Schools/districts need to provide time to allow teachers to meet repeatedly during the year.
Classroom-based tests (e.g., DIBELS and the Diagnostic Reading Assessment)	Provide training for elementary teachers in the appropriate use of these instruments, how often they should be used, and how to record results so that student growth across time points can be determined.	Classroom-based tests were designed primarily to help teachers track progress and adjust instruction accordingly. Because students differ in reading ability in early elementary grades and have a range of growth trajectories, it will be challenging to compare relative teachers’ contributions.
Student performance	Provide all art teachers in the district with the opportunity to meet and agree upon levels of performance (i.e., characteristics of a “beginning” performance and an “advanced” performance and how to document the performances to serve as evidence). The same applies to other classes for which a product or performance is the basis for the grade (e.g., music, drama, industrial arts classes).	If teachers do not have standards and a curriculum for the grade/subject, then they must first agree on what students should know and be able to do in a particular grade and subject before they can determine what different levels of performance should look like.
Other classroom-based evidence	Create opportunities for teachers in particular grades and subjects to meet together and agree upon ways to assess student learning. For example, timed multiplication drills might be used to document students’ growth in skills over time, but teachers must agree to a set of materials and a timeframe for conducting the drills.	Teacher-created assessments, worksheets, student journals, records of experiments, and other types of evidence can be excellent sources of documentation of student growth between two points in time, but there must be some consistency across classrooms and teachers to make such evidence comparable.

Whether utilizing existing measures, designing new ones, or using a combination of both, states and districts need to ensure that the measure or method utilized does not take time away from teaching. Instead, these assessments need to be an integral part of the teaching cycle that can quickly gauge student growth and inform teacher practice. Adding complicated, labor-intensive measures and processes will likely result in an upheaval from the education community and threaten the validity of the results.

Measures That May Improve Teacher Performance

All measures are not created equally in terms of how much they can inform a teacher about his or her practice and success in teaching specific content. Measures that are distant from the classroom, such as standardized tests administered once per year, are less likely to influence teaching practice and student learning in a timely manner, whereas measures that are aligned with an integral part of the curriculum and instructional sequence may provide useful information to the teacher about which skills and knowledge students have already mastered. This type of feedback, such as results from a pretest administered early in the year, can be used to guide instructional decisions.

In addition, ongoing assessments and examination of student work, especially in cooperation with colleagues, may not be included as part of teacher evaluation but may be useful for teachers in determining next steps for their students. When teachers know areas in which the students are experiencing difficulty, they can use that information to make the necessary instructional adjustments (e.g., reteaching), allowing extra opportunities for practice, instruction in small groups, peer tutoring, computer-assisted instruction, individual tutoring, or other changes in the method or type of instruction. In addition, teachers find value in working together to examine and score student work (e.g., essays, portfolios, or projects). Discussions with other

teachers about the differences between an outstanding piece of work and a good one can be valuable to teachers in thinking about how to target specific criteria in their own instruction.

Little attention has been paid to how the instruments and processes of teacher evaluation can inform professional growth opportunities. A feedback loop should be established that allows teachers and those who support them to identify areas of student weakness and strategize ways to improve instructional practices, resulting in improved student performance. Evaluation results should feed directly into that loop, providing specific, timely information in a format that is useful to teachers, administrators, and support personnel.

STATE GUIDANCE TO DISTRICTS

Districts will look to states for specific guidance about how to evaluate teachers' contributions to student learning growth, particularly in the nontested subjects and grades. There are several areas in which they need guidance.

Comparability: Across or Within Districts?

In order to better understand the differences among teacher effectiveness across schools and districts and identify teachers who are performing at high levels or those who are struggling, all teachers ideally would be evaluated in exactly the same way, using exactly the same measures. The state must first decide whether to insist on comparability *within* or *across* districts. A statewide system would be based on across-district comparability, whereas a district model would be based on within-district comparability. The following questions may be useful in making this decision:

- Is there a single set of subject-specific and grade-specific state standards for students that all districts use? If not, comparability across districts will be problematic.

- Do all districts throughout the state use the same curriculum and textbooks for all subjects? If not, it may be difficult to identify a common set of assessments that are appropriate for all districts.
- Do all districts have the same school calendar (e.g., start and end dates for the students, standardized testing dates, breaks, and holidays)? If not, it may be difficult to standardize the assessment process so that students are assessed at the same time across the state. The more standardized the assessment process is, the more comparable results will be.
- Do various types of educators in all districts across the state have the same job descriptions? The job description for some educators, particularly counselors, special educators, school nurses, librarians, and itinerant teachers, may vary widely from district to district.

If state staff answer “no” to any or all of these questions, they may want to consider comparability within rather than *across* districts. However, states could still provide guidelines to districts to ensure as much comparability as possible, given the district-to-district differences. For more information about appropriate guidance, see Goe, Holdheide, and Miller (in press).

Measures

States need to provide guidance to districts in selecting appropriate standards-based measures for documenting student growth. The following questions may help in determining the type of guidance to provide:

- Does the state want to approve all measures used by districts? If not, the state can provide the districts with guidelines and criteria for acceptable measures and leave approval of measures up to the districts.
- Does the state or district have a valid test that measures students’ progress toward mastery of grade-level and subject standards? If not, other measures will have

to be identified, purchased, or created to provide valid indicators of student growth. Districts can pool resources to share the costs of assessments and measures as a more cost-effective approach than each district attempting to pay these costs individually.

- Do districts have the capacity to implement processes for assessing student growth? If not, districts may need to join with other districts in regional or other purposeful consortiums to take advantage of economies of scale. For example, a number of rural districts might share information and resources, whereas an urban district might join forces with other urban districts in the state to form a consortium to share resources.

Exceptions

After a state or district adopts specific measures and processes for determining student learning growth, decision makers need to consider how to manage “exceptions” to the established processes for using these measures. For example, should a teacher be held accountable if the student was only assigned to his or her class for a portion of the school year? Or what happens if the student rarely attends school? Should the same level of accountability or attribution be assigned? Should working conditions be considered as a factor in determining teachers’ contributions to student learning growth? States and districts, working closely with teachers, administrators, and stakeholder groups, need to determine which exceptions to include and how to include them in ways that will ensure fairness and comparability.

Approaches to handling these exceptions may be left up to districts, but states may provide guidance or limit options to ensure greater comparability across districts.

Table 3. Priorities, Challenges, and Potential Solutions

<p>Measuring student growth between “two points in time”</p>	<p>Students complete only the pretest but not the posttest or vice versa.</p> <p>Students fail to turn in required work (e.g., a portfolio or project being used as the postmeasure).</p>	<p>With large numbers of students (e.g., at the secondary level), eliminate the student from the pool of students used to calculate the average student growth for the teacher.</p> <p>With smaller class sizes, it is important to include as many students as possible to reduce the margin of error. Allowing a review of other student work (homework or classwork), comparing current work or scores to those from previous years, or devising standards-based projects for students to complete are possible options, though imperfect at best.</p>
<p>Ensuring “rigor” of assessments</p>	<p>The measures used are complex, and it is difficult to determine rigor.</p> <p>There is little agreement about what rigor is and how it is reflected in the measures.</p>	<p>For a portfolio, project, or other multi-part measure, break down the components by the standard(s) being addressed. Will success on these components provide a clear indication of students’ mastery of standards-based knowledge or skills?</p> <p>Subject and grade-level standards should provide the focus for all measures. If the measure is not adequate to show progress toward mastery of standards-based skills and knowledge, it is not rigorous. In addition, demonstration of mastery of the knowledge and skills should be possible with the measure.</p>
<p>Making certain that measurement is “comparable across classrooms”</p>	<p>Raters are not adequately trained in scoring students’ work for portfolios, projects, performances, and products (the four Ps) that are being used as measures of students’ growth.</p> <p>Teachers acting as raters do not have time in their schedules to work with “like” teachers on scoring writing samples, portfolios, projects, performances, products (the four Ps), and so on.</p> <p>Pretests and posttests are not given in a standardized way.</p>	<p>Essays and the four Ps (i.e., portfolios, projects, performances, and products) all require training with scoring rubrics to ensure that all raters agree on what each level of the rubric looks like. Raters may be teachers, administrators, district personnel, or people hired specifically for scoring, but they must be trained to a high level of agreement. In addition, retraining and calibration should be conducted periodically to ensure that raters are still in agreement on interpreting the evidence. Training involves examining and discussing student work and rating it, then discussing rating decisions until agreement is reached.</p> <p>When teachers are trained as raters, it is important that they are given time to work together on scoring student work. Greater reliability and thus greater comparability will be achieved with multiple raters working together. Using some scheduled professional development time, grade-level or subject-level meeting time, or team time may be necessary.</p> <p>Results will not be comparable across classrooms unless specific practices are followed in giving pretests and posttests. These practices require a commitment and coordination across schools within a district to (1) choose a date/time that all schools agree to for pretesting of a subject/grade; (2) ensure that teachers are properly instructed on how to prepare students for the pretests and posttests; (3) give the tests at the same time of day; and (4) give tests for a predetermined length of time.</p>

Ongoing Research on Systems, Models, and Measures

Changes in teacher evaluation policies have occurred at a dizzying pace, outstripping researchers' ability to study the validity and fairness of the systems themselves and the individual components of the systems. Although research has been conducted on some of the measures, studies generally focus on low-stakes evaluation systems. (For a review of research on measures, see Goe, Bell, and Little, 2008.) There is little research on using student achievement growth as a measure of teacher effectiveness in a high-stakes system in which the results could mean commendation or probation, rewards or even dismissal. Planning for and consistently evaluating the relative quality of results from the use of various measures is important to increasing ability to accurately determine teacher effectiveness.

As states and districts implement evaluation systems that include multiple measures of student learning, it will be possible to evaluate the usefulness of various measures in differentiating among educators' levels of performance. This type of research should result in enhanced ability to conduct teacher evaluations that provide a nuanced, comprehensive, and accurate picture of teachers' contributions to student learning growth.

Considerations for States: Moving Forward

Without a research base to guide states' efforts, the TQ Center encourages caution and careful deliberation in designing and implementing high-stakes evaluation systems that measure teachers' contributions to

student learning growth. States may consider the following as they move forward:

- Partner with national and regional comprehensive centers in conducting needs assessments and outlining steps to take in determining appropriate measures and processes.
- Bring stakeholders (e.g., teachers, administrators, parents, school board members, union representatives, business leaders) to the table early in the discussions about measures and seek their help in communicating results.
- If the state does not currently have grade-level and subject standards for all courses, adopting such standards is important to ensure appropriate rigor in measuring student learning growth.
- The following steps can be used for selecting measures:
 - Categorize teachers by whether they are in tested or nontested subjects and grades.
 - Develop indicators within data systems to link teachers to appropriate student growth data.
 - Determine whether there are existing measures that might be useful in measuring student growth, and establish an approval process and/or listing of acceptable measures.
 - Secure content expertise to help evaluate coverage (i.e., whether measures exist to show learning growth for all teachers).
 - When gaps are found in existing measures, purchase or develop appropriate measures.
 - Consider alternative assessments as well as how measures need to be modified or differentiated through accommodations for students with special needs.

- Conserve resources by encouraging districts to join forces with other districts or regional groups to determine appropriate measures for nontested subjects and grades. This approach also contributes to greater comparability because teachers will be using the same measures across schools, districts, and regions.
- Consider whether human resources and capacity are sufficient to ensure fidelity of implementation.
- Develop a communication strategy to increase awareness and buy-in. Consider “frequently asked questions” pages on state and district websites and other means of sharing information about how and why measures were chosen and how they will be used.
- Establish a plan to evaluate measures to determine whether they can effectively differentiate among teacher performance.
- Evaluate processes and data each year and make needed adjustments.

CONCLUSION

There is little doubt that teacher evaluation has been permanently and irrevocably changed. No longer is a score on a principal’s observation checklist acceptable as evidence that a teacher is effective in the classroom. Linking teachers with student outcomes—including evidence of their growth in standards-based knowledge and skills—will become increasingly common. Moving forward in a responsible, deliberate, and cautious manner will ensure that the results are valid and defensible.

REFERENCES

- Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65–75.
- Battelle for Kids. (2009). *The importance of accurately linking instruction to students to determine teacher effectiveness*. Columbus, OH: Author. Retrieved February 18, 2011, from http://static.battelleforkids.org/images/BFK/Link_whitepagesApril2010web.pdf
- Betebenner, D. W. (2008). *A primer on student growth percentiles*. Dover, NH: National Center for the Improvement of Educational Assessment. Retrieved February 18, 2011, from <http://www.cde.state.co.us/cdedocs/Research/PDF/Aprimeronstudentgrowthpercentiles.pdf>
- Braun, H., Chudowsky, N., & Koenig, J. A. (2010). *Getting value out of value-added: Report of a workshop*. Washington, DC: National Academies Press. Retrieved February 18, 2011, from http://www.nap.edu/openbook.php?record_id=12820&page=1
- Feng, L., & Sass, T. R. (2009). *Special education teacher quality and student achievement*. Unpublished working paper.
- Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved February 18, 2011, from <http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf>
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved February 18, 2011, from <http://www.tqsource.org/publications/EvaluatingTeachEffectiveness.pdf>
- Goe, L., Holdheide, L., & Miller, T. C. (in press). *A practical guide to designing state teacher evaluation systems*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of early basic literacy skills* (6th ed.). Eugene, OR: Institute for the Development of Educational Achievement.
- Good, R. H., & Kaminski, R. A. (2011). *DIBELS next assessment manual*. Longmont, CO: Sopris.
- Herman, J. L., Heritage, M., & Goldschmidt, P. (in press). *Guidance for developing and selecting student growth measures for use in teacher evaluation*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Holdheide, L., Goe, L., Croft, A., & Reschly, D. (2010). *Challenges in evaluating special education teachers and English language learner specialists* (Research & Policy Brief). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved February 18, 2011, from <http://www.tqsource.org/publications/July2010Brief.pdf>
- Individuals with Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647 (2004). Retrieved February 18, 2011, from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=108_cong_public_laws&docid=f:publ446.108.pdf
- Koedel, C., & Betts, J. R. (2009). *Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein critique*. Cambridge, MA: National Bureau of Economic Research. Retrieved February 18, 2011, from http://economics.missouri.edu/working-papers/2009/WP0902_koedel.pdf
- McCaffrey, D., Sass, T. R., Lockwood, J. R., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, 4(4), 572–606. Retrieved February 18, 2011, from <http://www.mitpressjournals.org/doi/pdf/10.1162/edfp.2009.4.4.572>
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002). Retrieved February 18, 2011, from <http://www.ed.gov/policy/elsec/leg/esea02/index.html>

- Prince, C. D., Schuermann, P. J., Guthrie, J. W., Witham, P. J., Milanowski, A. T., & Thorn, C. A. (2009). *The other 69 percent: Fairly rewarding the performance of teachers of nontested subjects and grades*. Washington, DC: Center for Educator Compensation Reform. Retrieved February 18, 2011, from <http://www.cecr.ed.gov/guides/other69Percent.pdf>
- Reschly, D. R., & Holdheide, L. R. (2010, September 21). Figure presented during discussion group convened at "Evaluating and Rewarding Effectiveness and Navigating the Evolving Landscape," a symposium cohosted by the National Center on Performance Incentives, Battelle for Kids, and Vanderbilt Peabody College, Nashville, TN.
- Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417–458. Retrieved February 18, 2011, from <http://www.econ.ucsb.edu/~jon/Econ230C/HanushekRivkin.pdf>
- Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247–256. Retrieved February 18, 2011, http://www.sas.com/govedu/edu/ed_eval.pdf
- Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement* (No. R11-0435-02-001-97). Knoxville: University of Tennessee Value-Added Research and Assessment Center. Retrieved February 18, 2011, <http://www.mccsc.edu/~curriculum/cumulative%20and%20residual%20effects%20of%20teachers.pdf>
- Schochet, P. Z., & Chiang, H. S. (2010). *Error rates in measuring teacher and school performance based on student test score gains*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved February 18, 2011, from <http://ies.ed.gov/ncee/pubs/20104004/pdf/20104004.pdf>
- Secretary's Priorities for Discretionary Grant Programs, 75 Fed. Reg. 47,288 (proposed Aug. 5, 2010). Retrieved February 18, 2011, from <http://www2.ed.gov/legislation/FedRegister/other/2010-3/080510d.pdf>
- Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, 73(1), 89–122.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Brooklyn, NY: The New Teacher Project. Retrieved February 18, 2011, from <http://widgeteffect.org/downloads/TheWidgetEffect.pdf>

ABOUT THE NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

The National Comprehensive Center for Teacher Quality (TQ Center) was created to serve as the national resource to which the regional comprehensive centers, states, and other education stakeholders turn for strengthening the quality of teaching—especially in high-poverty, low-performing, and hard-to-staff schools—and for finding guidance in addressing specific needs, thereby ensuring that highly qualified teachers are serving students with special needs.

The TQ Center is funded by the U.S. Department of Education and is a collaborative effort of ETS, Learning Point Associates, and Vanderbilt University. Integral to the TQ Center's charge is the provision of timely and relevant resources to build the capacity of regional comprehensive centers and states to effectively implement state policy and practice by ensuring that all teachers meet the federal teacher requirements of the current provisions of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act.

The TQ Center is part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five content centers that provide expert assistance to benefit states and districts nationwide on key issues related to current provisions of ESEA.

IDAPA 08 - STATE BOARD OF AND STATE DEPARTMENT OF EDUCATION

08.02.02 - RULES GOVERNING UNIFORMITY

NOTICE OF RULEMAKING - PROPOSED RULE

THE FOLLOWING IS THE PROPOSED TEXT FOR SBOE REVIEW
-SCHEDULED FOR APRIL 2012-**120. LOCAL DISTRICT EVALUATION POLICY – TEACHER AND PUPIL PERSONNEL CERTIFICATE HOLDERS.**

Each school district board of trustees will develop and adopt policies for teacher performance evaluation using multiple measures in which criteria and procedures for the evaluation of certificated personnel are research based and aligned to Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. The process of developing criteria and procedures for certificated personnel evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the certificated personnel for whom it is written. (3-29-10)

01. Standards. Each district evaluation model shall be aligned to state minimum standards that are based on Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. Those domains and components include: (3-29-10)

- a. Domain 1 - Planning and Preparation: (3-29-10)
 - i. Demonstrating Knowledge of Content and Pedagogy; (3-29-10)
 - ii. Demonstrating Knowledge of Students; (3-29-10)
 - iii. Setting Instructional Goals Outcomes; (3-29-10)
 - iv. Demonstrating Knowledge of Resources; (3-29-10)
 - v. Designing Coherent Instruction; and (3-29-10)
 - vi. ~~Assessing~~ Designing Student Learning Assessments. (~~3-29-10~~)()
- b. Domain 2 - ~~Learning~~ The Classroom Environment: (~~3-29-10~~)()
 - i. Creating an Environment of Respect and Rapport; (3-29-10)
 - ii. Establishing a Culture for Learning; (3-29-10)
 - iii. Managing Classroom Procedures; (3-29-10)
 - iv. Managing Student Behavior; and (3-29-10)
 - v. Organizing Physical Space. (3-29-10)
- c. Domain 3 - Instruction and Use of Assessment: (3-29-10)
 - i. Communicating ~~Clearly and Accurately~~ with Students; (~~3-29-10~~)()
 - ii. Using Questioning and Discussion Techniques; (3-29-10)
 - iii. Engaging Students in Learning; (3-29-10)
 - iv. ~~Providing Feedback to Students~~ Using Assessment in Instruction; and (~~3-29-10~~)()

- v. Demonstrating Flexibility and Responsiveness; ~~and~~ (3-29-10)()
- vi. ~~Use Assessment to Inform Instruction and Improve Student Achievement.~~ (3-29-10)
- d. Domain 4 - Professional Responsibilities: (3-29-10)
 - i. Reflecting on Teaching; (3-29-10)
 - ii. Maintaining Accurate Records; (3-29-10)
 - iii. Communicating with Families; (3-29-10)
 - iv. ~~Contributing to the School and District~~ Participating in a Professional Community; (3-29-10)()
 - v. Growing and Developing Professionally; and (3-29-10)
 - vi. Showing Professionalism. (3-29-10)

02. Parent Input. For evaluations conducted on or after July 1, 2012, input from the parents and guardians of students shall be considered as a factor in the evaluation of any school-based certificated employees. For such certificated employees on a Category A, B or grandfathered renewable contract, this input shall be part of the first half of the evaluation that must be completed before February 1 of each year (Section 33-513 and 33-514, Idaho Code). ()

03. Student Achievement. For evaluations conducted on or after July 1, 2012, all certificated employees must receive an evaluation in which at least fifty percent (50%) of the evaluation results are based on objective measures of growth in student achievement as determined by the board of trustees and based upon research. This student achievement portion of the evaluation shall be completed by the end of the school year in which the evaluation takes place (Section 33-513 and 33-514, Idaho Code). ()

04. Participants. Each district evaluation policy will include provisions for evaluating all certificated employees identified in Section 33-1001, Idaho Code, Subsection 136, and each school nurse and librarian (~~Section 33-515, Idaho Code~~). Policies for evaluating certificated employees should identify the differences, if any, in the conduct of evaluations for nonrenewable contract personnel and renewable contract personnel. (4-1-97)()

05. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: (4-1-97)

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions. (4-1-97)

b. Evaluation criteria -- statements of the general criteria upon which certificated personnel will be evaluated. (4-1-97)

c. Evaluator -- identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation and after September 1, 2014, shall have proof of proficiency in evaluating teacher performance. (4-1-97)

d. Sources of data -- description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation should be included as one (1) source of data. (4-1-97)

e. Procedure -- description of the procedure used in the conduct of certificated personnel evaluations. (4-1-97)

f. Communication of results -- the method by which certificated personnel are informed of the results of evaluation. (4-1-97)

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel. (4-1-97)

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations. (4-1-97)

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action. (4-1-97)

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's personnel evaluation system. (4-1-97)

k. Professional development and training -- a plan for ongoing training for evaluators/administrators and teachers on the districts evaluation standards, tool and process. (3-29-10)

l. Funding -- a plan for funding ongoing training and professional development for administrators in evaluation. (3-29-10)

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. Aggregate data shall be the basis for the district's Needs Assessment in determining district-wide professional development. Individual performance data shall be the foundation of individualized Professional Performance Plans for all teachers. Professional Performance Plans shall be used in annual evaluation as a means of measuring professional growth. District shall implement use of Professional Growth Plans no later than January 1, 2015.

n. A plan for how evaluations will be used to identify proficiency and define a process that identifies and assists teachers in need of improvement. No later than March 01, 2014, districts shall have established an individualized teacher evaluation rating system with a ranking of not proficient, basic, proficient, and distinguished. Districts shall ensure that an Individualized Professional Development plan is created for each teacher based upon evaluation findings, and to be used in subsequent years as the baseline measurement for professional development and growth.

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and administrators in the development and ongoing review of their teacher evaluation plan. (3-29-10)

046. Evaluation Policy - Frequency of Evaluation. The evaluation policy should include a provision for evaluating all certificated personnel on a fair and consistent basis. *At a minimum, the policy must provide standards for evaluating the following personnel: All contract personnel shall be evaluated at least once annually.* (4-1-97)()

a- First, second, and third year nonrenewable contract personnel will be evaluated at least once prior to the beginning of the second semester of the school year. (4-1-97)

b- All renewable contract personnel will be evaluated at least once annually. (4-1-97)

057. Evaluation Policy - Personnel Records. Permanent records of each certificated personnel evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code). (4-1-97)

08. Evaluation System Approval. Each school district board of trustees will develop and adopt policies for teacher and pupil personnel certificated performance evaluation in which criteria and procedures for the evaluation of are research based. Once developed, each district shall submit the system of evaluation to the State Department of Education for approval prior to formal adoption. By January 1, 2014 an evaluation plan which incorporates all of the above elements shall be submitted to the State Department of Education for approval. Once approved, subsequent changes made in the evaluation system shall be resubmitted for approval.

121. LOCAL DISTRICT EVALUATION POLICY - ADMINISTRATIVE CERTIFICATE HOLDERS.

Each school district board of trustees will develop and adopt policies for administrator performance evaluation in which criteria and procedures for the evaluation of administratively certificated personnel are research based. The process of developing criteria and procedures for certificated personnel evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the certificated personnel for whom it is written.

01. Standards. Each district evaluation model shall be aligned to state minimum standards, including proof of proficiency in conducting teacher evaluations using the state's adopted model, the Charlotte Danielson Framework for Teaching. Proof of proficiency in evaluating teacher performance shall be required of all administrators. **Attachment 23, Page 3 of 5**

02. Parent Input. For evaluations conducted on or after July 1, 2012, input from the parents and guardians of students shall be considered as a factor in the evaluation of any administratively certificated employees and must be completed before February 1 of each year (Section 33-513 and 33-514, Idaho Code).

03. Student Achievement. For evaluations conducted on or after July 1, 2012, all administratively certificated employees must receive an evaluation in which at least fifty percent (50%) of the evaluation results are based on objective measures of growth in student achievement as determined by the board of trustees and based upon research. This student achievement portion of the evaluation shall be completed by the end of the school year in which the evaluation takes place (Section 33-513 and 33-514, Idaho Code).

04. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: (4-1-97)

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions. (4-1-97)

b. Evaluation criteria -- statements of the general criteria upon which administratively certificated personnel will be evaluated.

c. Evaluator -- identification of the individuals responsible for appraising or evaluating administratively certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation.

d. Sources of data -- description of the sources of data used in conducting administratively certificated personnel evaluations. Proficiency in conducting evaluations through classroom observation should be included as one (1) source of data.

e. Procedure -- description of the procedure used in the conduct of administratively certificated personnel evaluations.

f. Communication of results -- the method by which administratively certificated personnel are informed of the results of evaluation

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change.

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations.

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's personnel evaluation system

k. Professional development and training -- a plan for ongoing training for evaluators/administrators and teachers on the districts evaluation standards, tool and process.

l. Funding -- a plan for funding ongoing training and professional development for administrators in evaluation.

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. Aggregate data shall be the basis for the district's Needs Assessment in determining district-wide professional development for administrators. Individual performance data shall be the foundation of individualized Professional Performance Plans. Professional Performance Plans shall be used in annual evaluation as a means of measuring professional growth in instructional leadership. District shall implement use of Professional Growth Plans no later than January 1, 2015.

n. A plan for how evaluations will be used to identify proficiency and define a process that identifies and assists administrative personnel in need of improvement. No later than March 01, 2014, districts shall have established an individualized evaluation rating system with a ranking of not proficient, basic, proficient, and distinguished. Districts shall ensure that an Individualized Professional Development plan is created for each administrative certificate holder based upon evaluation findings, and to be used in subsequent years as the baseline measurement for professional development and growth.

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and administrators in the development and ongoing review of their teacher evaluation plan.

05. Evaluation Policy - Frequency of Evaluation. The evaluation policy should include a provision for evaluating all certificated personnel on a fair and consistent basis. All contract personnel shall be evaluated at least once annually.

06. Evaluation Policy - Personnel Records. Permanent records of each certificated personnel evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code).

07. Evaluation System Approval. Each school district board of trustees will develop and adopt policies for teacher and pupil personnel certificated performance evaluation in which criteria and procedures for the evaluation of are research based. Once developed, each district shall submit the system of evaluation to the State Department of Education for approval prior to formal adoption. . By January 1, 2014 an evaluation plan which incorporates all of the above elements shall be submitted to the State Department of Education for approval. Once approved, subsequent changes made in the evaluation system shall be resubmitted for approval.

School District:

Date:



**TEACHER PERFORMANCE EVALUATION
STATE STANDARDS AND REQUIREMENTS RUBRIC**

The districts teacher evaluation model is based on or is aligned to the following minimum standards:	Met	Partially Met	Not Met	Comments:
Domain 1 – Planning and Preparation 1a: Demonstrating Knowledge of Content and Pedagogy				
Domain 1 – Planning and Preparation 1b: Demonstrating Knowledge of Students				
Domain 1 – Planning and Preparation 1c: Setting Instructional Goals				
Domain 1 – Planning and Preparation 1d: Demonstrating Knowledge of Resources				
Domain 1 – Planning and Preparation 1e: Designing Coherent Instruction				
Domain 1 – Planning and Preparation 1f: Assessing Student Learning				
Domain 2 – Learning Environment 2a: Creating an Environment of Respect and Rapport				
Domain 2 – Learning Environment 2b: Establishing a Culture for Learning				
Domain 2 – Learning Environment 2c: Managing Classroom Procedures				
Domain 2 – Learning Environment 2d. Managing Student Behavior				



The districts teacher evaluation policy includes the following provisions:	Met	Partially Met	Not Met	Comments:
District evaluation policy includes a provision for evaluating all certificated employees identified in Section 33-1001, Idaho Code, Subsection 13, and each school nurse and librarian (Section 33-515, Idaho Code). Policies for evaluating certificated employees should identify the differences, if any, in the conduct of evaluations for nonrenewable contract personnel and renewable contract personnel.				
District evaluation policy contains statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions.				
District evaluation policy contains statements of the general criteria upon which certificated personnel will be evaluated.				
District evaluation policy contains identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation.				
District evaluation policy contains description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation should be included as one (1) source of data.				
District evaluation policy contains description of the procedure used in the conduct of certificated personnel evaluations.				
District evaluation policy contains the method by which certificated personnel are informed of the results of evaluation.				
District evaluation policy contains the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel.				
District evaluation policy contains the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations.				
District evaluation policy contains the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action.				
District evaluation policy contains a description of the method used to monitor and evaluate the district's personnel evaluation system.				

<p>Districts evaluation policy includes a plan for including all stakeholders, including teachers, school board members and administrators, in the development and ongoing review of the teacher evaluation plan.</p>				
<p>District evaluation policy contains a plan for how evaluations will be used to identify proficiency and define a process that identifies and assists teachers in need of improvement</p>				
<p>District evaluation policy contains a plan for ongoing training and professional development for evaluators/administrators and teachers on the district’s evaluation standards, tool and process.</p>				
<p>District evaluation policy contains a plan for funding ongoing training and professional development for administrators in evaluation</p>				
<p>District evaluation policy contains a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development</p>				
<p>District evaluation policy contains at a minimum, a provision for evaluating the following personnel:</p> <ul style="list-style-type: none"> • First-, second-, and third-year nonrenewable contract personnel will be evaluated at least once prior to the beginning of the second semester of the school year. • All renewable contract personnel will be evaluated at least once annually. 				
<p>Permanent records of each certificated personnel evaluation will be maintained in the employee’s personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy.</p>				

State of Idaho Teacher Performance Evaluation Implementation Guidelines

Every Teacher Performance Evaluation Model Must Include the Following:

Performance Levels: Each district must identify descriptors of performance levels for each domain, which will, at a minimum, address proficient and unsatisfactory practice. Example of performance levels a district might identify include: unsatisfactory, basic, proficient, distinguished. In recognition of research into mastery, proficient performance in a domain is meeting 80% of the components.

Reliability and Validity: Part of the vision of the Teacher Performance Evaluation Task Force is for each district's evaluation tool and process to be valid and reliable and utilize data to support those qualifications. Districts will report content validity data within the first year - gather input from those being evaluated on the indicators within components and domains (this meets the requirements in the Idaho Administrative Code 08.02.02.120). Reliability is demonstrated through the plan for ongoing training for evaluators to ensure that different evaluators recognize the same behaviors at the same level of performance.

Training and Professional Development: As part of each district's process and implementation of a teacher evaluation model, there must be a plan for ongoing training for evaluators/administrators as well as professional development for teachers on the district's evaluation tool and process. Districts must ensure that all administrators responsible for performing evaluations be trained in the district approved evaluation model.

Required Components of a District Teacher Evaluation Model:

- Districts must adopt or develop a teacher evaluation model that is aligned to state minimum standards that are based on Charlotte Danielson's Framework for Teaching domains and components of instruction.
- Districts will develop or adopt their own instruments and procedures for evaluating teachers based on these standards.
- The evaluation process will be determined by the local district providing that it meets the minimum number of evaluations per year required in Idaho laws and rules.
- Each district's teacher evaluation model must include, at a minimum, the following information:

- **Purpose** -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions.
- **Evaluation criteria** -- statements of the general criteria upon which certificated personnel will be evaluated.
- **Evaluator** -- identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation.
- **Sources of data** -- description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation should be included as one (1) source of data.
- **Procedure** -- description of the procedure used in the conduct of certificated personnel evaluations.
- **Communication of results** -- the method by which certificated personnel are informed of the results of evaluation.
- **Personnel actions** -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel.
- **Appeal** -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations.
- **Remediation** -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action.
- **Monitoring and evaluation** -- A description of the method used to monitor and evaluate the district's personnel evaluation system.
- **Funding** -- a plan for funding ongoing training and professional development for administrators in evaluation.
- **Collecting and using data** -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development.
- **Identify proficiency** -- A plan for how evaluations will be used to identify proficiency and define a process that identifies and assists teachers in need of improvement.
- **Stakeholders** -- A plan for including all stakeholder including, but not limited to, teachers, board members and administrators in the development and ongoing review of their teacher evaluation plan.
- **Professional development and training** -- a plan for ongoing training for evaluators/administrators and teachers on the districts evaluation standards, tool and process.

- The task force believes that reliability is developed and demonstrated through ongoing training for evaluators.
- Districts must ensure that all administrators responsible for performing evaluations be trained in the district's state-approved evaluation model.
- Districts must identify what funds they are currently utilizing for administrator professional development in evaluation as well as funds they will utilize to support ongoing training and professional development.

State Approval:

Every school district and charter school must submit its evaluation model to the State Department of Education for approval by February 2010.

To be approved, the evaluation model must meet the minimum statewide standards for teacher evaluations and the minimum number of evaluations per year as required by Idaho laws and rules. Models must also address performance levels, reliability and validity, and ongoing training and professional development. A team of reviewers at the State Department of Education who are trained in the framework will approve the evaluation models.

Plans that are not approved will be returned to the districts highlighting recommendations for change. The State Department of Education will establish a process of appeals for districts that wish to contest a plan that was not approved.

IDAPA 08 - STATE BOARD OF AND STATE DEPARTMENT OF EDUCATION

08.02.02 - RULES GOVERNING UNIFORMITY

DOCKET NO. 08-0202-1106

NOTICE OF RULEMAKING - PROPOSED RULE

THE FOLLOWING IS THE PROPOSED TEXT OF DOCKET NO. 08-0202-1106

120. LOCAL DISTRICT EVALUATION POLICY.

Each school district board of trustees will develop and adopt policies for teacher performance evaluation in which criteria and procedures for the evaluation of certificated personnel are research based and aligned to Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. The process of developing criteria and procedures for certificated personnel evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the certificated personnel for whom it is written. (3-29-10)

01. Standards. Each district evaluation model shall be aligned to state minimum standards that are based on Charlotte Danielson's Framework for Teaching Second Edition domains and components of instruction. Those domains and components include: (3-29-10)

- a. Domain 1 - Planning and Preparation: (3-29-10)
 - i. Demonstrating Knowledge of Content and Pedagogy; (3-29-10)
 - ii. Demonstrating Knowledge of Students; (3-29-10)
 - iii. Setting Instructional Goals Outcomes; ~~(3-29-10)~~()
 - iv. Demonstrating Knowledge of Resources; (3-29-10)
 - v. Designing Coherent Instruction; and (3-29-10)
 - vi. ~~Assessing~~ Designing Student Learning Assessments. ~~(3-29-10)~~()
- b. Domain 2 - ~~Learning~~ The Classroom Environment: ~~(3-29-10)~~()
 - i. Creating an Environment of Respect and Rapport; (3-29-10)
 - ii. Establishing a Culture for Learning; (3-29-10)
 - iii. Managing Classroom Procedures; (3-29-10)
 - iv. Managing Student Behavior; and (3-29-10)
 - v. Organizing Physical Space. (3-29-10)
- c. Domain 3 - Instruction and Use of Assessment: (3-29-10)
 - i. Communicating ~~Clearly and Accurately~~ with Students; ~~(3-29-10)~~()
 - ii. Using Questioning and Discussion Techniques; (3-29-10)
 - iii. Engaging Students in Learning; (3-29-10)
 - iv. ~~Providing Feedback to Students~~ Using Assessment in Instruction; and ~~(3-29-10)~~()

- v. Demonstrating Flexibility and Responsiveness; ~~and.~~ (3-29-10)()
- vi. ~~Use Assessment to Inform Instruction and Improve Student Achievement.~~ (3-29-10)
- d. Domain 4 - Professional Responsibilities: (3-29-10)
 - i. Reflecting on Teaching; (3-29-10)
 - ii. Maintaining Accurate Records; (3-29-10)
 - iii. Communicating with Families; (3-29-10)
 - iv. ~~Contributing to the School and District~~ Participating in a Professional Community; (3-29-10)()
 - v. Growing and Developing Professionally; and (3-29-10)
 - vi. Showing Professionalism. (3-29-10)

02. Parent Input. For evaluations conducted on or after July 1, 2012, input from the parents and guardians of students shall be considered as a factor in the evaluation of any school-based certificated employees. For such certificated employees on a Category A, B or grandfathered renewable contract, this input shall be part of the first half of the evaluation that must be completed before February 1 of each year (Section 33-513 and 33-514, Idaho Code). ()

03. Student Achievement. For evaluations conducted on or after July 1, 2012, all certificated employees must receive an evaluation in which at least fifty percent (50%) of the evaluation results are based on objective measures of growth in student achievement as determined by the board of trustees. This student achievement portion of the evaluation shall be completed by the end of the school year in which the evaluation takes place (Section 33-513 and 33-514, Idaho Code). ()

024. Participants. Each district evaluation policy will include provisions for evaluating all certificated employees identified in Section 33-1001, Idaho Code, Subsection 1-36, and each school nurse and librarian (~~Section 33-515, Idaho Code~~). Policies for evaluating certificated employees should identify the differences, if any, in the conduct of evaluations for nonrenewable contract personnel and renewable contract personnel. (4-1-97)()

035. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: (4-1-97)

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions. (4-1-97)

b. Evaluation criteria -- statements of the general criteria upon which certificated personnel will be evaluated. (4-1-97)

c. Evaluator -- identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should have received training in evaluation. (4-1-97)

d. Sources of data -- description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation should be included as one (1) source of data. (4-1-97)

e. Procedure -- description of the procedure used in the conduct of certificated personnel evaluations. (4-1-97)

f. Communication of results -- the method by which certificated personnel are informed of the results of evaluation. (4-1-97)

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel. (4-1-97)

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations. (4-1-97)

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action. (4-1-97)

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's personnel evaluation system. (4-1-97)

k. Professional development and training -- a plan for ongoing training for evaluators/administrators and teachers on the districts evaluation standards, tool and process. (3-29-10)

l. Funding -- a plan for funding ongoing training and professional development for administrators in evaluation. (3-29-10)

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. (3-29-10)

n. A plan for how evaluations will be used to identify proficiency and define a process that identifies and assists teachers in need of improvement. (3-29-10)

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and administrators in the development and ongoing review of their teacher evaluation plan. (3-29-10)

046. Evaluation Policy - Frequency of Evaluation. The evaluation policy should include a provision for evaluating all certificated personnel on a fair and consistent basis. ~~At a minimum, the policy must provide standards for evaluating the following personnel:~~ All contract personnel shall be evaluated at least once annually. (4-1-97)()

~~**a.** First, second, and third year nonrenewable contract personnel will be evaluated at least once prior to the beginning of the second semester of the school year.~~ (4-1-97)

~~**b.** All renewable contract personnel will be evaluated at least once annually.~~ (4-1-97)

057. Evaluation Policy - Personnel Records. Permanent records of each certificated personnel evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code). (4-1-97)

teachscape
*Framework for Learning
Proficiency System*



Rating Protocol: Classroom Practice Observation Form

This agreement with the observer is a part of the assessment process. The observer will observe the classroom practice and provide a rating based on the Framework for Teaching. The observer will provide a rating for each component of the Framework for Teaching. The observer will provide a rating for each component of the Framework for Teaching. The observer will provide a rating for each component of the Framework for Teaching.

The observer will provide a rating for each component of the Framework for Teaching. The observer will provide a rating for each component of the Framework for Teaching. The observer will provide a rating for each component of the Framework for Teaching.

- Identify and record instructional evidence
- Align teaching evidence to a specific component of the Framework for Teaching
- Evaluate evidence to render accurate observations about classroom practice in the various components of the Framework for Teaching
- Score classroom practice based on the evidence

Subject: **MATH** Grade: **5** UID: **00000000000000000000000000000000** Score: **3** Item ID: **00000000000000000000000000000000**

Score Feedback

Component	Your Score	Actual Score
Managing Student Behavior	4	2
Engaging Students in Learning	3	5
Using Assessment in Instruction	2	3

Managing Student Behavior
Your score is too high. Numerous students are inattentive and talking in a disruptive manner. The teacher is inconsistent in monitoring and implementing the standards of conduct. Teacher asks students to put phones away, but doesn't address off-topic conversation.

Engaging Students in Learning
Your score is correct. All students are cognitively engaged with the tasks and activities. The students are working independently with some teacher instruction and clarification, but all students are expected to solve and complete their own work. The content is challenging and developmentally appropriate to the students and allow students to develop and construct knowledge (after solving for area, students are required to compare measurement benchmarks to see which items best fit the solution). The pacing of the lesson seems appropriate for all students.

Using Assessment in Instruction
Your score is too low. The teacher does use assessment regularly during instruction. Teacher gives specific feedback that advances learning. The teacher sees that students are struggling with one of the problems. Teacher says "Area equals length times width. So, think about what 2 numbers would multiply by one another to equal that area, okay?"

Proficiency Test: Elementary and Middle School Testing

The Elementary and Middle School Testing (EMST) is a computer-based assessment that measures student proficiency in English language arts, math, and science or social studies content areas.

The EMST is a computer-based assessment that measures student proficiency in English language arts, math, and science or social studies content areas. The EMST is a computer-based assessment that measures student proficiency in English language arts, math, and science or social studies content areas.



- Five test versions: elementary, middle school, high school, K-8, and K-12
- Includes English language arts, math, and science or social studies content areas
- Takes about 7.5 hours and contains two pass/not pass stages
- Includes test passage rate reports for observers and proficiency analysis reports for district administrators

Framework for Teaching Proficiency System

877.98.TEACH



The Framework for Teaching Proficiency System (FTPS) was developed by the ETS, Inc. and the California State Board of Education (SBE) in partnership with the California State Board of Education (SBE) and the California State Board of Education (SBE). It is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California. The FTPS is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California.

Implemented Successfully in the MET Project

The Framework for Teaching Proficiency System (FTPS) was implemented successfully in the MET Project. The MET Project is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California. The FTPS is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California.

Developed with Charlotte Danielson and ETS

The Framework for Teaching Proficiency System (FTPS) was developed with Charlotte Danielson and ETS. The FTPS is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California. The FTPS is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California.



“The Framework for Teaching Proficiency System is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California. It is a comprehensive, research-based framework for teaching proficiency that is designed to be used by all teachers in California.”

—Charlotte Danielson

877.98.TEACH
info@teachscope.com
71 Stevenson St., 5th Floor
San Francisco, CA 94105
www.teachscope.com

Great teachers create inspired learners.

To find out how the Teachscope Framework for Teaching Proficiency System can help you, call 877.98.TEACH, or visit our website at www.teachscope.com.

Christina P. Linder

From: Christina P. Linder
Sent: Wednesday, December 28, 2011 12:33 PM
To: 'adunn@sugarsalem.com'; 'coburnr@d93.k12.id.us'; 'rcampbell@sd60.k12.id.us'; 'bjolley@sd60.k12.id.us'; 'gmlowe@wendellschools.org'; 'jshawver@kimberly.edu'; 'ward@nsd131.org'; 'jrapp@lewistonschools.net'; 'CoburnR@d93.k12.id.us'; 'mgreen@falsonridgecharter.org'
Cc: Teresa Burgess; Becky Martin
Subject: Proficiency for Evaluators Using the Danielson Framework - Regional Face to Face Danielson Training for Administrators and Teacher Leaders
Attachments: Framework for Teaching Proficiency System.pdf

Dear District Administrators,

I am writing to let you know that we've been able to confirm dates for our regional face to face trainings. Our goal is to provide you with the opportunity to participate in trainings that will lead to inter-rater reliability as you strive to effectively implement the Danielson Framework within your district.

You will not only have the opportunity to send up to 20 administrators and teacher leaders to be trained, but each will be able to become officially certified as an evaluator. The brochure that explains this process is attached above. Through a combination of face-to-face trainings and the availability of video training and practice tests, your evaluators will be able to prove that they have achieved proficiency in evaluating teachers for both formative and summative purposes.

The state will provide training, materials, meals, and reimbursement for up to five substitute teachers to facilitate participation of teacher leaders. We would ask that you consider using Title IIA funds to pay for travel expenses and perhaps provide a stipend for participants. Completion of this "train-the-trainer" training will result in an opportunity for you to build capacity within your own districts, and provide ongoing support for administrators and teachers. The dates are outlined in the table below.

Next week you will receive a second email with a link to register for the trainings. Please forward the email with the embedded link to all administrators and teachers leaders who will be participating. This link will not only have a brief survey to give our trainers and idea of the level of expertise among participants, but will also be used as the official vehicle for confirming registration of participants.

Proficiency for Evaluators Using the Danielson Framework - Training Schedule 2011-2012

Participants: Principals, Evaluators and Teacher Leaders

Region 1 and 2 CdA or Post Falls Facility to be arranged Time: 8:30am-3:30pm	Region 3 Boise or Meridian Time: 8:30am-3:30pm	Region 4 ,5 and 6 Idaho Falls Facility to be arranged Time: 8:30am-3:30pm
Day 1- October 19,2011	Day 1-January 18, 2012	Day 1- January 20,2012
Day 2 -February 28, 2012	Day 2- March 8, 2012	Day2-March 6, 2012
Day 3-April 17,2012	Day 3-April 19, 2012	Day 3-April 24, 2012
Day 4 – TBA : Next Steps	Day 4- June 14, 2012	Day 4- June 7, 2012

Certification: Certified Evaluator	Certification: Certified Evaluator	Certification: Certified Evaluator
------------------------------------	------------------------------------	------------------------------------

Please don't hesitate to call or email any of us listed above if you have questions. I hope you were able to get some rest and peace over the long break, and have come back feeling refreshed. Our team so looks forward to working with you.

Warm regards,
c

Christina Linder

Director
Certification and Professional Standards
Idaho State Department of Education
(208) 332-6886

WISE Tool - School Turnaround Plan Scoring Rubric

Note to Reviewers: Meaningful interventions designed to improve the academic achievement of students in One Star Turnaround Plan Schools (i.e., priority schools) must be aligned with all of the following “turnaround principles”. In the first year of review, objectives and plans must be created that align with the turnaround principles, do not score the monitoring and implementation elements. In the second year and beyond, score the objectives, tasks, and monitoring and implementation elements. A plan must be marked as acceptable or exceptional in all categories in order to be approved.

Principle 1 - Provide strong leadership by: (1) reviewing the performance of the current principal; (2) either replacing the principal if such a change is necessary to ensure strong and effective leadership, or demonstrating to the SEA that the current principal has a track record in improving achievement and has the ability to lead the turnaround effort; and (3) providing the principal with operational flexibility in the areas of scheduling, staff, curriculum, and budget.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 2 - Ensure that teachers are effective and able to improve instruction by: (1) reviewing the quality of all staff and retaining only those who are determined to be effective and have the ability to be successful in the turnaround effort; (2) preventing ineffective teachers from transferring to these schools; and (3) providing job-embedded, ongoing professional development informed by the teacher evaluation and support systems and tied to teacher and student needs.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 3 - Redesign the school day, week, or year to include additional time for student learning and teacher collaboration.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 4 - Strengthen the school's instructional program based on student needs and ensuring that the instructional program is research-based, rigorous, and aligned with State academic content standards.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 5 - Use data to inform instruction and for continuous improvement, including by providing time for collaboration on the use of data.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 6 - Establish a school environment that improves school safety and discipline and addressing other non-academic factors that impact student achievement, such as students' social, emotional, and health needs.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Principle 7 - Provide ongoing mechanisms for family and community engagement.

PLANNING ELEMENTS	NEEDS REVISION	ACCEPTABLE	EXCEPTIONAL
<i>Objectives Created</i>	The objectives planned are not specific, measureable, or attainable; or, the objectives do not align with the required turnaround principle.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and they are likely to result in academic improvement.	The planned objectives are specific, measureable, and attainable. They align with the required turnaround principle, and are likely to result in dramatic improvement that is systemic and sustainable.
<i>Tasks Developed</i>	Created tasks are not evident, not sufficient to accomplish the objectives, not aligned to the objective, or not realistic.	Created tasks align to the objectives created, demonstrate a specific set of steps to accomplish the plan, and are likely to bring about the intended improvement in the school.	Created tasks align to the objectives created, represent a clear and concise focus on improvement, demonstrate a specific set of steps to accomplish the plan, and are likely to cause dramatic and sustained improvement.
<i>Monitoring and Implementation (applicable after implementation year 1)</i>	The plan has not been monitored, implemented, and/or evidence for the implementation of planned objectives is insufficient.	The plan is being monitored and evidence is provided that objectives and tasks are being implemented as intended.	The plan is being monitored, timely adjustments are taking place when presented with obstacles, and evidence of improvement is presented both in the plan and during Focus Visits.

Idaho Department of Education
July 11, 2012

EMPHASIS ON GROWTH

Idaho was asked by the U.S. Department of Education to answer the following questions.

- A.** Please address concerns regarding Idaho's proposed accountability system:
- Address concerns regarding the over-reliance on student growth in the overall accountability system (*e.g.*, growth to achievement and growth to achievement subgroup measures combine for 50% of the overall score for high schools and 75% of the overall score for middle and elementary schools). *See 2.A.i.a.*
 - Please explain Idaho's rationale for the Student Growth Percentile (SGP) ranges used in the Adequate Growth Flowchart and address the concern that schools that did not meet Adequate Growth Percentile (AGP) are still able to achieve the maximum score with a minimally greater SGP than schools that meet the AGP. Also, demonstrate or explain any safeguards in place to ensure that a school cannot score well on the growth index while not helping students make sufficient progress toward proficiency. *See 2.A.i.a.*

STUDENT GROWTH

Within the Idaho system, achievement is represented in several ways. First, the percentage of proficient students is measured under the “Achievement” Category for three subject areas. This traditional metric, the same one used to calculate proficiency in the Adequate Yearly Progress system, accounts for 25% of the overall rating for elementary schools. In schools with grade 12 (i.e. High Schools), it accounts for 20% of the overall rating. Achievement is also measured in High Schools under the “Postsecondary and Career Readiness” category. Under this metric, which accounts for 30% of the overall rating, achievement of 11th graders on a college entrance or placement exam and participation and achievement in advanced opportunities for 11th and 12th graders, account for half of these points. In high schools, achievement therefore accounts for 35% of the overall rating.

Idaho also uses the same Student Growth Percentile (SGP) and Adequate Growth Percentile (AGP) model as Colorado. As is consistent with Colorado’s already approved system, the AGP calculations, the criterion-referenced indicator measuring whether students are making enough growth to reach proficiency, are actually an additional achievement measure.

Table 24 in Section 2.B.i. has been reformatted to illustrate the percentage of schools achieving points through either meeting AGP or not meeting AGP.

Table 1
Growth to Achievement Point Distribution
Percentage of Schools (2010-11)

Points	Subject	Met AGP		Points	Subject	Did not meet AGP	
	<i>Median Student Growth Percentile</i>	Schools			<i>Median Student Growth Percentile</i>	Schools	
	Reading	(N=576)	%		Reading	(N=8)	%
5	66-99	13	2.3	5	70-99	-	0
4	52-65	225	39.1	4	61-69	-	0
3	43-51	266	46.2	3	51-60	-	0
2	30-42	72	12.5	2	36-50	1	12.5
1	1-29	-	0	1	1-35	7	87.5
	Mathematics	(N=525)	%		Mathematics	(N=58)	%
5	66-99	41	7.8	5	70-99	-	0
4	52-65	216	41.1	4	61-69	-	0
3	43-51	189	36	3	51-60	1	1.7
2	30-42	79	15	2	36-50	26	44.8
1	1-29	-	0	1	1-35	31	53.4
	Language Usage	(N=525)	%		Language Usage	(N=55)	%
5	66-99	20	3.81	5	70-99	-	0
4	52-65	217	41.3	4	61-69	-	0
3	43-51	239	45.5	3	51-60	1	1.8
2	30-42	49	9.3	2	36-50	30	54.5
1	1-29	-	0	1	1-35	24	43.6

Table 1 illustrates two things clearly: 1) the majority of Idaho schools have met their AGP goals, and 2) for those schools that have not met the growth targets of AGP, none are given higher than 3 points within the metric. In fact, greater than 85% of the schools that fail to meet these growth targets receive only 2 points or less. For example, within the “Growth to Achievement” category if a school does not meet AGP and does not have at least a median SGP of 51st percentile, the highest points awarded would be only 40% of the total points or 20 out of 50 for an elementary school and 12 out of 30 for schools

with a grade 12. This same calculation is used for the “Growth to Achievement Subgroups” and therefore if we assume that the subgroups performed as well as the overall population the highest total points awarded would be 10 out of 25 for elementary schools and 8 out of 20 for schools with grade 12. Table 2 below uses these calculations for a school without a grade 12 illustrating that even if the school receives all possible points in the Achievement category (which is highly unlikely given the performance described in the other metrics) the highest rating this school could receive is a Three Star.

Table 2
Example Overall Rating Chart for School without Grade 12

Accountability Measures	Points Achieved	Points Eligible	Star Rating
Achievement	25	25	
Growth to Achievement	20	50	
Growth to Achievement Gaps	10	25	
TOTAL	55	100	***
Participation Rates	Were at least 95% of students Tested?	Yes	***
STAR RATING	Three Star		

STUDENT GROWTH AND ADEQUATE GROWTH PERCENTILES

Idaho replicated some of the same analyses conducted by Colorado using the same procedures with Idaho’s growth data. Idaho’s data follows a similar pattern to Colorado’s.

In the first analysis, the median SGPs is aggregated by school. In Figure 1, the distribution of median SGPs in Mathematics for schools without a grade 12 forms a bell curve with a clustering around the 50th percentile, as we would expect. There are very few schools in the extremes in the 80th percentile and in the 20th percentile. Figure 2 is the same data represented through a cumulative density function in Mathematics. The function uses the same data to illustrate the cumulative percentage of schools at specific median student growth percentiles.

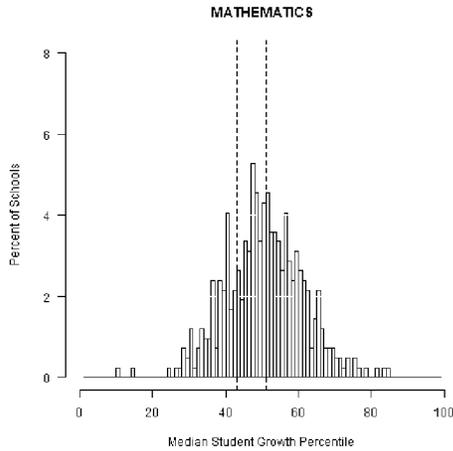


Figure 1. Distribution of schools without grade 12 mathematics median SGP

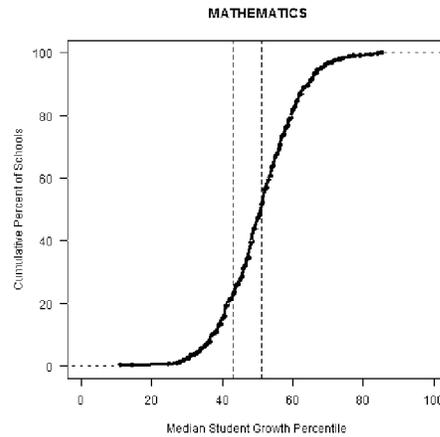
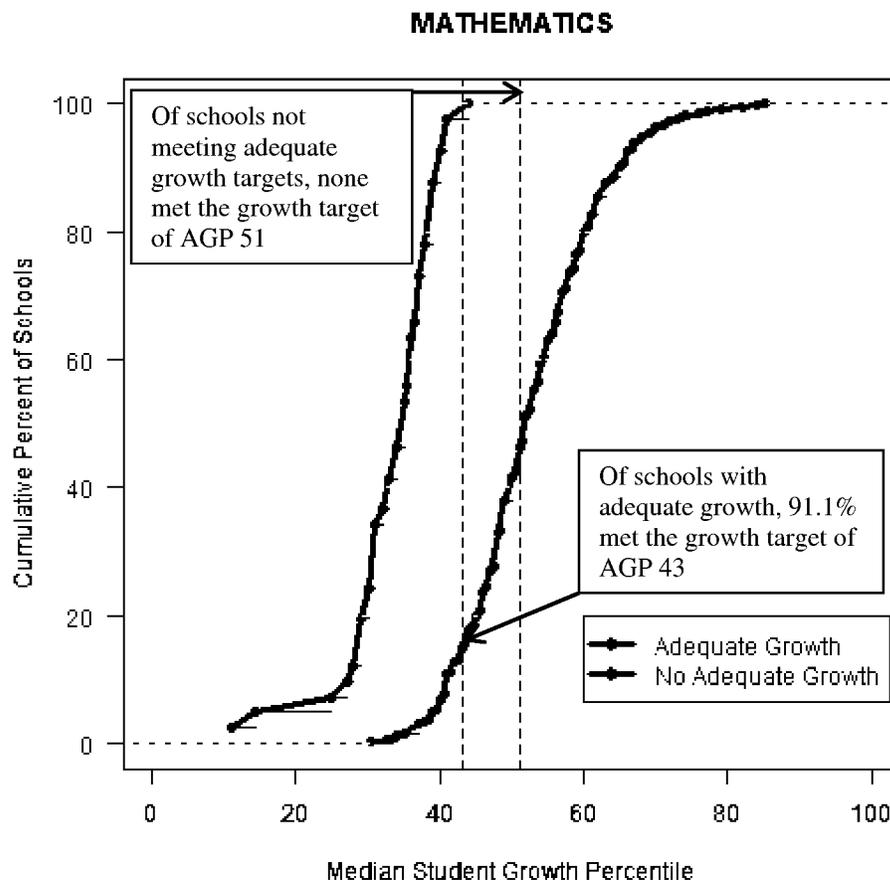


Figure 2. Cumulative density function of schools without grade 12 mathematics median SGP



**Figure 3. Cumulative density function of schools without grade 12 mathematics:
Comparison of schools making AGP and not making AGP**

Figure 3 shows the pattern of median SGPs for schools without a grade 12 making growth targets (AGP) and those that do not. The red line illustrates those schools that did not meet the growth targets. The vertical dotted lines illustrate threshold for 3 points: 51 for schools not making adequate growth in red and 43 for schools making adequate growth in blue. These schools that did not meet the growth targets show a much lower median SGP overall. Approximately 22% of schools without a grade 12 have a median SGP below the 43rd percentile. In addition, 43% of schools without a grade 12 have a median SGP below the 51st percentile. For schools making AGP, 91.1% of these schools met or exceeded the threshold for 3 points (43). The percentages of schools at each of these points are important because they are set at the lowest median SGP a school can achieve (both through making AGP and not making AGP targets, respectively) and receive at least 3 points on the 5 point matrix. While the median SGP of 43 and 51 may seem low, compared to the actual performance within Idaho shows that 44% of Idaho schools are achieving 4 or 5 points on the 5 point matrix. Further, for schools to receive 4 points on the 5 point matrix and yet failed to meet achieve AGP goals, they must have a median SGP of 61st percentile. Referring back to Table 1 on page 2, we can see that no Idaho schools that failed to make AGP goals received 4 points for SGP in any subject area.

Table 3
**Correlation Matrix Showing Relationships among Current Year Proficiency,
Median SGP and Median AGP
School without Grade 12**

Subject	Achievement to Median SGP	Achievement to Median AGP	Median SGP to Median AGP
Reading	0.46	-0.74	-0.30
Mathematics	0.43	-0.76	-0.13
Language Usage	0.50	-0.80	-0.26

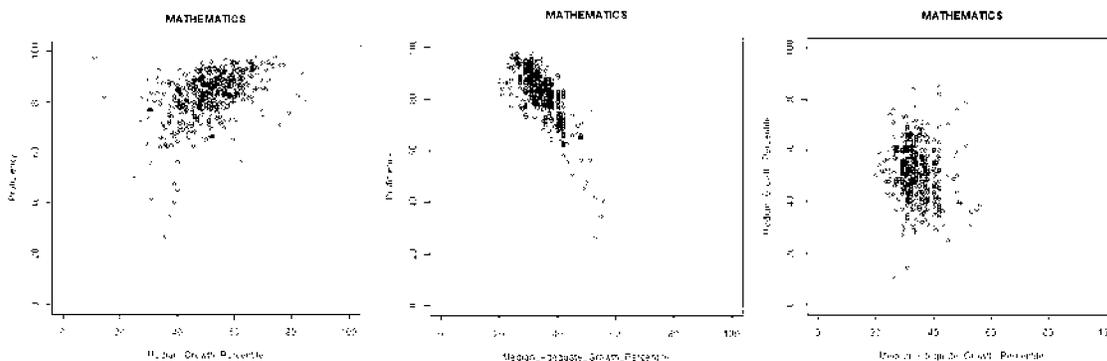


Figure 4. Scatterplot comparing schools without grade 12 median AGP and proficiency (left), proficiency and median AGP (center), and median AGP and median SGP for mathematics.

Table 3 shows the relationships of schools without a grade 12 regarding 1) comparing median SGP and proficiency, 2) median AGP and proficiency and, 3) median AGP and median SGP. Figure 4 illustrates these relationships in scatterplots. The first chart shows a weak relationship between median SGP and proficiency. This is an expected relationship and is the same as was found in Colorado. SGP measures the amount of student growth and proficiency measures reaching a specific bar. The second chart shows a negative correlation between proficiency and AGP. This again is consistent with Colorado’s data and is consistent with the design of the model. As stated by Colorado: “Students starting out from low score have to grow more, and students already achieving high test scores need to grow less” to reach proficiency. The third chart is a relationship much like that in first chart. There is a weak relationship between median SGP and AGP. This chart illustrates the differences between the normative SGP and criterion-referenced AGP.

CONCLUSION

Idaho built its accountability system elements around Growth to Achievement and Growth to Achievement Subgroups being highly informed by the current and the several previous years of research done by Colorado in implementing the SGP and AGPs. Given that Idaho has only had the Student Growth Percentiles model in place for a little over a year, it was a benefit to be able to analyze the results of a fairly developed system and inform the setting of goals based on that research in tandem with Idaho’s one year performance. Granted, the overall Idaho system is unique in many ways to the one used in Colorado, but the procedures and formulas underlying the growth calculations are the same. Further, as illustrated by the correlations, scatterplots, cumulative density functions and the AGP tables, Idaho’s model is equally as rigorous. The goals set are ambitious yet achievable and strongly account for continued achievement for a necessary sustained growth.

**IDAPA 08 - STATE BOARD OF AND STATE DEPARTMENT OF
EDUCATION**

08.02.02 – RULES GOVERNING

UNIFORMITY

**NOTICE OF RULEMAKING – PROPOSED
RULE**

**120. LOCAL DISTRICT EVALUATION POLICY – TEACHER AND PUPIL PERSONNEL
CERTIFICATE HOLDERS.**

Each school district board of trustees will develop and adopt policies for teacher performance evaluation using multiple measures in which criteria and procedures for the evaluation of certificated personnel are research based and aligned to Charlotte Danielson’s Framework for Teaching Second Edition domains and components of instruction. The process of developing criteria and procedures for certificated personnel evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the certificated personnel for whom it is written. (3-29-10)

01. Standards. Each district evaluation model shall be aligned to state minimum standards that are based on Charlotte Danielson’s Framework for Teaching Second Edition domains and components of instruction. Those domains and components include: (3-29-10)

- a.** Domain 1 - Planning and Preparation: (3-29-10)
 - i. Demonstrating Knowledge of Content and Pedagogy; (3-29-10)
 - ii. Demonstrating Knowledge of Students; (3-29-10)
 - iii. Setting Instructional Goals-Outcomes; (3-29-12)
 - iv. Demonstrating Knowledge of Resources; (3-29-10)
 - v. Designing Coherent Instruction; and (3-29-10)
 - vi. Designing Student Assessments. (3-29-12)
- b.** Domain 2 - The Classroom Environment: (3-29-12)
 - i. Creating an Environment of Respect and Rapport; (3-29-10)
 - ii. Establishing a Culture for Learning; (3-29-10)
 - iii. Managing Classroom Procedures; (3-29-10)
 - iv. Managing Student Behavior; and (3-29-10)
 - v. Organizing Physical Space. (3-29-10)
- c.** Domain 3 - Instruction and Use of Assessment: (3-29-10)
 - i. Communicating with Students; (3-29-12)

- ii. Using Questioning and Discussion Techniques; (3-29-10)
- iii. Engaging Students in Learning; (3-29-10)
- iv. Using Assessment in Instruction; and (3-29-12)
- v. Demonstrating Flexibility and Responsiveness. (3-29-12)
- d. Domain 4 - Professional Responsibilities:** (3-29-10)
 - i. Reflecting on Teaching; (3-29-10)
 - ii. Maintaining Accurate Records; (3-29-10)
 - iii. Communicating with Families; (3-29-10)
 - iv. Participating in a Professional Community; (3-29-12)
 - v. Growing and Developing Professionally; and (3-29-10)
 - vi. Showing Professionalism. (3-29-10)

02. Parent Input. For evaluations conducted on or after July 1, 2011, input from the parents and/or guardians of students shall be considered as a factor in the evaluation of any school-based certificated employees. For such certificated employees on a Category A, B or grandfathered renewable contract, this input shall be part of the first portion of the evaluation (as stipulated in 33-514(4), Idaho Code,) that must be completed before February March 1 of each year (Section 33-513 and 33-514, Idaho Code). (3-29-12)

03. Student Achievement. For evaluations conducted on or after July 1, 2012, all certificated employees must receive an evaluation in which at least fifty percent (50%) of the evaluation results are based on objective measures of growth in student achievement as determined by the board of trustees and based upon research.. This student achievement portion of the evaluation shall be completed by the end of the school year in which the evaluation takes place (Section 33-513 and 33-514, Idaho Code), and must include growth in student achievement as measured by the Idaho Student Achievement Test (ISAT). (3-29-12)

04. Participants. Each district evaluation policy will include provisions for evaluating all certificated employees identified in Section 33-1001, Idaho Code, Subsection 16, and each school nurse and librarian. Evaluations shall be differentiated for pupil personnel certificate holders in a way that aligns with the Framework for Teaching to what extent possible. Policies for evaluating certificated employees should identify the differences, if any, in the conduct of evaluations for certificated employees on a Category A, B or grandfathered renewable contract nonrenewable contract personnel and renewable contract personnel. (3-29-12)

05. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: (4-1-97)

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional improvement, personnel decisions. (4-1-97)

b. Evaluation criteria -- statements of the general criteria upon which certificated personnel will be evaluated. (4-1-97)

c. Evaluator -- identification of the individuals responsible for appraising or evaluating certificated personnel performance. The individuals assigned this responsibility should shall have received training in evaluation and after September 1, 2014, shall have proof of proficiency in evaluating teacher performance. (4-1-97)

d. Sources of data -- description of the sources of data used in conducting certificated personnel evaluations. For classroom teaching personnel, classroom observation ~~should~~shall be included as one (1) source of data. (4-1-97)

e. Procedure -- description of the procedure used in the conduct of certificated personnel evaluations. (4-1-97)

f. Communication of results -- the method by which certificated personnel are informed of the results of evaluation. (4-1-97)

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. Note: in the event the action taken as a result of evaluation is to not renew an individual's contract or to renew an individual's contract at a reduced rate, school districts should take proper steps to follow the procedures outlined in Sections 33-513 through 33-515, Idaho Code in order to assure the due process rights of all personnel. (4-1-97)

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of certificated personnel evaluations. (4-1-97)

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action. (4-1-97)

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's personnel evaluation system. (4-1-97)

k. Professional development and training -- a plan for ongoing training and professional learning based upon the district's evaluation standards for evaluators/administrators and teachers on the districts evaluation standards, tool and process. (3-29-10)

l. Funding -- a plan for funding ongoing training and professional development for administrators in evaluation. (3-29-10)

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. Aggregate data shall be the basis for the district's Needs Assessment in determining district-wide professional development. Individual performance data shall be the foundation of an Individual Professional Learning Plan for all teachers. Individual Professional Learning Plans shall be used in the annual evaluation as a means of measuring professional growth. Districts shall implement the use of Professional Growth Plans no later than September 1, 2014. (3-29-10)

n. Individualizing teacher evaluation rating system -- A plan for how evaluations will be used to identify proficiency and record growth over time. No later than March 01, 2014, districts shall have established an individualized teacher evaluation rating system with a ranking of unsatisfactory being equal to "1", basic being equal to "2", proficient being equal to "3", and distinguished being equal to "4" . Districts shall ensure that an Individualized Professional Learning Plan is created for each teacher based upon evaluation findings, and shall be used in subsequent years as the baseline measurement for professional development and growth define a process that identifies and assists teachers in need of improvement. (3-29-10)

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and administrators in the development and ongoing review of their teacher evaluation plan. (3-29-10)

06. Evaluation Policy - Frequency of Evaluation. The evaluation policy shall include a provision for evaluating all certificated personnel on a fair and consistent basis. All contract personnel shall be evaluated at least once annually. An annual evaluation (Section 33-514 and 33-515) shall include, at a minimum, two (2)

formative observations and/or evaluative discussions.

(3-29-12)

07. Evaluation Policy - Personnel Records. Permanent records of each certificated personnel evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code). (4-1-97)

08. Evaluation System Approval. Each school district board of trustees will develop and adopt policies for teacher and pupil personnel certificated performance evaluation in which criteria and procedures for the evaluation of are research based and aligned with the Framework for Teaching. Once developed, each district shall submit the system of evaluation to the State Department of Education for approval prior to formal adoption. By January 1, 2014 an evaluation plan which incorporates all of the above elements shall be submitted to the State Department of Education for approval. Once approved, subsequent changes made in the evaluation system shall be resubmitted for approval. ()

121. LOCAL DISTRICT EVALUATION POLICY - ADMINISTRATIVE CERTIFICATE HOLDERS.

Each school district board of trustees will develop and adopt policies for administrator performance evaluation in which criteria and procedures for the evaluation of administratively certificated personnel are research based. The process of developing criteria and procedures for administrator evaluation will allow opportunities for input from those affected by the evaluation; i.e., trustees, administrators and teachers. The evaluation policy will be a matter of public record and communicated to the administrator for whom it is written. ()

01. Standards. Each district administrator evaluation model shall be aligned to state minimum standards, including proof of proficiency in conducting teacher evaluations using the state's adopted model, the Charlotte Danielson Framework for Teaching. Proof of proficiency in evaluating teacher performance shall be required of all administrators no later than September 1, 2014. Administrator evaluation standards shall additionally address the following domains and components: ()

a. Domain 1: School Climate - An educational leader promotes the success of all students by advocating, nurturing and sustaining a school culture and instructional program conducive to student learning and staff professional development. An educational leader articulates and promotes high expectations for teaching and learning. ()

i. School Culture - Administrator establishes a safe, collaborative, and supportive culture ensuring all students are successfully prepared to meet the requirements for tomorrow's careers and life endeavors. ()

ii. Communication - Administrator is proactive in communicating the vision and goals of the school or district, the plans for the future, and the successes and challenges to all stakeholders. ()

iii. Advocacy - Administrator advocates for education, the district and school, teachers, parents, and students that engenders school support and involvement. ()

b. Domain 2: Collaborative Leadership - An educational leader promotes the success of all students by ensuring management of the organization, operations and resources for a safe, efficient and effective learning environment. In collaboration with others, uses appropriate data to establish rigorous, concrete goals in the context of student achievement and instructional programs. He/She uses research and/or best practices in improving the education program. ()

i. Shared Leadership - Administrator fosters shared leadership that takes advantage of individual expertise, strengths, and talents, and cultivates professional growth. ()

ii. Priority Management - Administrator organizes time and delegates responsibilities to balance administrative/managerial, educational, and community leadership priorities. ()

iii. Transparency - Administrator seeks input from stakeholders and takes all perspectives into consideration when making decisions. ()

iv. Leadership Renewal - Administrator strives to continuously improve leadership skills through professional development, self-reflection, and utilization of input from others. ()

v. Accountability - Administrator establishes high standards for professional, legal, ethical, and fiscal accountability self and others. ()

c. Domain 3: Instructional Leadership - An educational leader promotes the success of all students by facilitating the development, articulation, implementation, and stewardship of a vision of learning that is shared and supported by the school community. He/She provides leadership for major initiatives and change efforts and uses research and/or best practices in improving the education program. ()

i. Innovation - Administrator seeks and implements innovative and effective solutions that comply with general and special education law. ()

ii. Instructional Vision - Administrator insures that instruction is guided by a shared, research-based instructional vision that articulates what students do to effectively learn the subject. ()

iii. High Expectations - Administrator sets high expectation for all students academically, behaviorally, and in all aspects of student well-being. ()

iv. Continuous Improvement of Instruction - Administrator has proof of proficiency in assessing teacher performance based upon the Danielson Framework for Teaching. Aligns resources, policies, and procedures toward continuous improvement of instructional practice guided by the instructional vision. ()

v. Evaluation- Administrator uses teacher/administrator evaluation and other formative feedback mechanisms to continuously improve teacher/administrator effectiveness. ()

vi. Recruitment and Retention -Administrator recruits and maintains a high quality staff. ()

02. Parent Input. For evaluations conducted on or after July 1, 2012, input from the parents and/ or guardians of students shall be considered as a factor in the evaluation of any school-based certificated employees (as stipulated in 33- 514(4), Idaho Code). ()

03. Student Achievement. For evaluations conducted on or after July 1, 2012, all certificated employees must receive an evaluation in which at least fifty percent (50%) of the evaluation results are based on objective measures of growth in student achievement as determined by the board of trustees and based upon research. This student achievement portion of the evaluation shall be completed by the end of the school year in which the evaluation takes place (Section 33-513 and 33-514, Idaho Code) and must include growth in student achievement as measured by the Idaho Student Achievement Test (ISAT). ()

04. Evaluation Policy - Content. Local school district policies will include, at a minimum, the following information: ()

a. Purpose -- statements that identify the purpose or purposes for which the evaluation is being conducted; e.g., individual instructional leadership, personnel decisions. ()

b. Evaluation criteria -- statements of the general criteria upon which administratively certificated personnel will be evaluated. ()

c. Evaluator -- identification of the individuals responsible for appraising or evaluating administratively certificated personnel performance. The individuals assigned this responsibility shall have received training in evaluation. ()

d. Sources of data -- description of the sources of data used in conducting administratively certificated personnel evaluations. Proficiency in conducting observations and evaluating effective performance shall be included as one (1) source of data. ()

e. Procedure -- description of the procedure used in the conduct of administratively certificated personnel evaluations. ()

f. Communication of results -- the method by which administratively certificated personnel are informed of the results of evaluation. ()

g. Personnel actions -- the action, if any, available to the school district as a result of the evaluation and the procedures for implementing these actions; e.g., job status change. ()

h. Appeal -- the procedure available to the individual for appeal or rebuttal when disagreement exists regarding the results of administrator evaluations. ()

i. Remediation -- the procedure available to provide remediation in those instances where remediation is determined to be an appropriate course of action. ()

j. Monitoring and evaluation. -- A description of the method used to monitor and evaluate the district's administrator evaluation system. ()

k. Professional development and training -- a plan for ongoing training and professional learning based upon the district's evaluation standards and process. ()

l. Funding -- a plan for funding ongoing training and professional development for evaluators of administrators. ()

m. Collecting and using data -- a plan for collecting and using data gathered from the evaluation tool that will be used to inform professional development. Individual performance data shall be the foundation of an Individual Professional Learning Plan for all administrators. Individual Professional Learning Plans shall be used in the annual evaluation as a means of measuring professional growth with an emphasis on instructional leadership. Districts shall implement the use of Professional Growth Plans no later than September 1, 2014. ()

n. Individualizing administrator evaluation rating system -- a plan for how evaluations will be used to identify proficiency and record growth over time. No later than March 01, 2014, districts shall have established an individualized administrator evaluation rating system with a ranking of unsatisfactory being equal to "1", basic being equal to "2", proficient being equal to "3", and distinguished being equal to "4". Districts shall ensure that an Individualized Professional Learning Plan is created for each administrator based upon evaluation findings, and shall be used in subsequent years as the baseline measurement for professional development and growth. ()

o. A plan for including all stakeholders including, but not limited to, teachers, board members, and administrators in the development and ongoing review of their administrator evaluation plan. ()

05. Evaluation Policy - Frequency of Evaluation. The evaluation policy should include a provision for evaluating all administrative personnel on a fair and consistent basis. An annual evaluation (Section 33-514 and 33-515) shall include, at a minimum, two (2) formative observations and evaluative discussions. ()

06. Evaluation Policy - Personnel Records. Permanent records of each administrator evaluation will be maintained in the employee's personnel file. All evaluation records will be kept confidential within the parameters identified in federal and state regulations regarding the right to privacy (Section 33-518, Idaho Code). ()

07. Evaluation System Approval. Each school district board of trustees will develop and adopt policies for administrator performance evaluation in which criteria and procedures for the evaluation are research based and aligned with state standards. Once developed, each district shall submit the system of evaluation to the

State Department of Education for approval prior to formal adoption. By January 1, 2014 an evaluation plan which incorporates all of the above elements shall be submitted to the State Department of Education for approval. Once approved, subsequent changes made in the evaluation system shall be resubmitted for approval. ()