AIR®

AMERICAN INSTITUTES FOR RESEARCH®

JUNE 2013

Washington Teacher/Principal Evaluation Project (Year 3) Report

Final Project Report to the Office of State Superintendent of Public Instruction

Jenni Fetters

Ellen Behrstock-Sherratt, Ph.D.

Bo Zhu

Washington Teacher/Principal Evaluation Project (Year 3) Report

Final Report to the Office of State Superintendent of Public Instruction

June 2013

Jenni Fetters

Ellen Behrstock-Sherratt, Ph.D

Bo Zhu

1120 East Diehl Road, Suite 200
Naperville, IL 60563-1486
800-356-2735 | 630-649-6500

www.air.org

Contents

Executive Summary

Washington State, like many other states across the country, is engaged in a comprehensive overhaul of our teacher and principal evaluation systems. Educator effectiveness continues to be a focal point and critical ingredient in the education reform dialogue. To this end, the Washington State Legislature passed ESSB 5895, which became state law on June 7, 2012. OPSI and the TPEP Steering Committee organizations guided the 2012-13 development work of the districts in Washington State as they begin implementation in September, 2013.

Background

The Teacher and Principal Evaluation Project (TPEP), which was created in Engrossed Second Substitute Senate Bill 6696 (E2SSB 6696) in the 2010 legislative session, offers Washington the opportunity to identify the measures of effective teaching and leading. Complimentary legislation, Engrossed Substitute Senate Bill 5895 (ESSB 5895), adds specificity to the statutes put forth in E2SSB 6696`. The new evaluation system must hold educators accountable and serve to leverage authentic professional growth. This emerging system was built on the foundation of the new teacher and principal evaluation criteria and developed by Washington educators. It provides a direction that will empower teachers, principals, and district leaders to meet the needs of all students in Washington. The new evaluation system sets high expectations for what teachers and principals should know and be able to do, values diversity, and fosters a high commitment to teaching and leading as professional practice.

TPEP Steering Committee

The TPEP work is led by a state-level steering committee comprised of the Office of Superintendent of Public Instruction (OSPI), Washington Education Association (WEA), Association of Washington School Principals (AWSP), Washington Association of School Administrators (WASA), Washington State Parent Teacher Association (WSPTA), and the Washington State School Directors' Association (WSSDA). These six organizations model the collaborative process required to implement this new evaluation system.

Instructional and Leadership Frameworks

Superintendent Dorn has approved the following instructional frameworks for use in the Washington State Teacher Professional Growth and Evaluation System:
Instructional
- The Center for Educational Leadership's (CEL) *5D+ Teacher Evaluation Rubric 2.0*
- Charlotte Danielson's *The Framework for Teaching (2011)*
- Robert Marzano's *The Marzano Teacher Evaluation Model*

Leadership
- The Association of Washington School Principal's (AWSP) *The AWSP Leadership Framework*
- Robert Marzano's *The Marzano School Leadership Evaluation Model*

For more information on the approved instructional and leadership frameworks and their role in the new evaluation system, visit http://tpep-wa.org/the-model/framework-and-rubrics/. To see the following map and charts in full size, visit http://tpep-wa.org/wp-content/uploads/Statewide framework overview.png.

Statewide Framework Selections
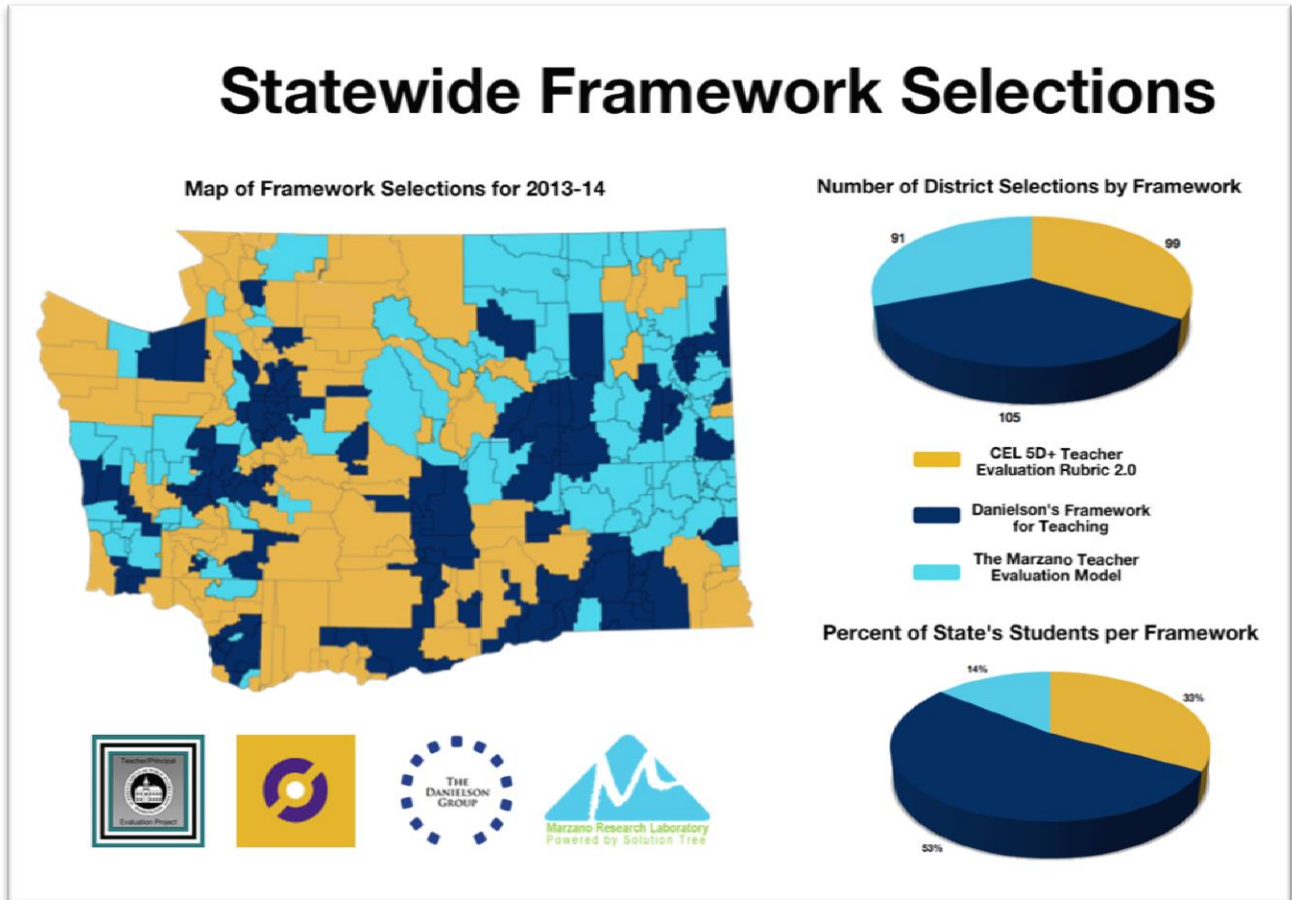
Regional Implementation Grants (RIGs)

During the 2012-13 school year, 208 districts applied and received grants to develop their educator evaluation models using the regionally based Educational Service Districts (ESDs). These districts used the learning from the first year of the TPEP pilot and, through a common curriculum formed the foundation of a comprehensive evaluation system for both teachers and principals.

OSPI will issue a 2013–14 grant application for Student Growth Regional Implementation Grants. These grants will be used to support the remaining districts to design, implement, and provide professional learning on the student growth portion of the new evaluation systems. The RIGs will use consistent curriculum designed by OSPI, the ESDs, and the TPEP steering committee organizations that will include:

- Planning for Student Growth Goal (SGG) Setting

- Identifying Tools/Measures for Student Growth Goals

- Setting Targets for Student Growth Goals

- Monitoring Student Growth Goals

- Reflecting and Evaluating Student Growth Goals

TPEP Status	Number of Districts
Pilot	16
RIG I	70
RIG II	108
SIG Districts (non-RIG only)	5
Districts piloting prior to 2013-14	229

Modules

In compliance with Washington State law and designed to assist Washington districts in designing and implementing each aspect of a comprehensive educator evaluation system that meets the requirements of SB 5895, the online professional learning module series is available for download by Educational Service Districts (ESDs) or district leaders. These modules will be used by the ESDs and district teams to establish the student growth portion of their systems using appropriated funding for the 2013-14 school year.

- *Introduction to Educator Evaluation in Washington.* The basics of educator evaluation reforms, the evaluation criteria for teachers and principals, the four-level rating system, state and local decision matrix and a preview of the remaining modules.

- *Using Instructional and Leadership Frameworks in Educator Evaluation.* An orientation to the components of instructional and leadership frameworks, how they are different from previous evaluation tools, and how they support identification of practice across a continuum. This module will provide a "jumpstart" into the three instructional and two leadership frameworks.

- *Preparing and Applying Formative Multiple Measures of Performance: An Introduction to Self-Assessment, Goal Setting, and Criterion Scoring.* An overview of the types of measures required and supported by RCW 28A.405.100 and WAC 392-191A, the differences between measures and evidence, how to move beyond an observation-only evaluation system, and the benefits to the

types of measures used in educator evaluation. This module includes criterion scoring guidance that has been informed by the instructional and leadership framework authors.

- *Including Student Growth in Educator Evaluation.* An overview of goal setting for student growth, selecting classroom-based, school-based, district-based, and state-based tools, and using student learning data in educator evaluation. This module will offer a process for establishing student growth goals, examples of student growth goals, and a process for determining the change in student achievement between two points in time.

- *Conducting High-Quality Observations and Maximizing Rater Agreement.* An overview of high-quality observation practices with special emphasis on collecting evidence, strategies for maximizing rater agreement, and strategies for districts and school administrators to learn about and plan for maximizing rater agreement.

- *Providing High-Quality Feedback for Continuous Professional Growth and Development.* An overview of examples of, and protocols for, how to provide feedback to teachers and leaders so that they continue to grow and improve in their practice; how to engage faculty in these conversations; and strategies for connecting professional development planning with evaluation outcomes.

- *Combining Multiple Measures Into a Summative Rating.* While an evaluation rating is often viewed as a measure of a single point in time, it is actually the culmination of a comprehensive process of self-assessment, goal-setting, plan implementation, dialogue, and reflection that unfolds over months. This module provides an understanding of how to assess practice using multiple types of evidence and performance rubrics in a thoughtful, comprehensive, reliable manner and follow the process from OSPI to create an overall performance rating.

State of the State Educator Survey: This statewide survey provided information on the progress of all Washington districts in preparing to implement educator evaluation reforms and garnered the perspectives of teachers, principals, superintendents, and school directors about the reforms. The survey results helped to inform OSPI's planning around ongoing supports that Washington districts need to implement new educator evaluation systems effectively (aligned with Project Objectives #1 and #2).

Policy-to-Practice Cohort Meetings: The original nine pilot sites were each assigned to a policy-to-practice cohort to focus on one of the key issues identified in SB 5895. AIR supported OSPI in facilitating three cohort meetings during Year 3, providing cohort participants with access to emerging research and practices in other states, and collecting data from the cohorts to share key insights and begin formulating recommendations on each topic area (aligned with Project Objective #2).

The following report describes each deliverable listed above, including a summary of the deliverable and an overview of the key findings or product associated with it. The report begins with an introduction to Washington's educator evaluation reforms and AIR's past work supporting OSPI and the TPEP Steering Committee from 2010 through 2013. After summarizing the key findings from the state of the state educator survey, the report provides an overview of the three policy-to-practice cohorts, their activities, and the recommendations and considerations generated under each issue area. The section concludes by reporting the results of a policy-to-practice cohort survey. The survey gathered participants' views on 35 recommendations the cohort formulated on using evaluation results for a range of human resource decisions. Finally, the report concludes by describing the communications and professional learning module deliverables.

Introduction

In 2010 the state of Washington initiated a set of legislative and policy reforms to improve educator evaluation in schools across the state. A core part of Washington's approach to this reform effort is the Teacher/Principal Evaluation Pilot (TPEP) project, which established nine educator evaluation pilot districts. [1] Starting in 2010, the Washington Office of Superintendent of Public Instruction (OSPI) contracted with the American Institutes for Research to provide research and technical assistance support to OSPI, the TPEP Steering Committee, and the nine pilot districts.

In Years 1 and 2 (2010–2012), AIR conducted a statewide survey, supported regional forms and focus groups, and conducted ongoing case studies of each pilot site and created a cross-case analysis integrating insights and key findings across all nine pilots. As a result of this work, OSPI and the TPEP Steering Committee have been able to make decisions about teacher and principal evaluation that are based both on the latest research emerging from around the country and on what their constituents are thinking and doing at the local level.

In Year 3 (2012–2013), AIR continued to support OSPI and the TPEP steering committee through systematic survey data collection and analysis, ongoing expert consultation, communications materials, professional learning modules, meeting facilitation, and comprehensive reporting on teacher and principal evaluation reform based on the experiences of pilot districts. These deliverables, and the key findings or products associated with them, are described in further detail in the sections that follow.

Legislative Background

In 2010 and 2012, the Washington State Legislature passed two significant legislative bills that have shaped educator evaluation reform in the state. Senate Bill 6696, passed in 2010, established the broad outlines of a new state policy on educator evaluation including requirements for multiple measures and new teaching and leadership criteria, as well as requiring the creation of the TPEP project to inform state-level policy making.

In a path-breaking approach to educator evaluation policy development, SB 6696 mandated the creation of nine pilot sites, and each site was provided grant funding for two years to support the design and implementation of a new teacher and principal evaluation system that fit the revised legislative requirements. The dedicated work of these nine pilots yielded experiences and lessons that continued to inform OSPI's ongoing development of policy and guidance on educator evaluation in Year 3 as they addressed additional changes in response to Senate Bill 5895.

In 2012, the legislature passed Senate Bill 5895, which elaborated additional requirements around the inclusion of measures of student growth and the use of three "preferred" instructional and leadership frameworks. In addition, the bill mandated that the TPEP Steering Committee and

---

[1] The pilot sites included the following districts: Anacortes, Central Valley, Kennewick, North Mason, North Thurston, Othello, Snohomish, Wenatchee, and a consortium of rural districts that included participants from Almira, Davenport, Liberty, Medical Lake, Pullman, Reardan-Edwall, Wellpinit, and Wilbur school districts.

nine pilot sites "continue to examine implementation issues and refine tools" for educator evaluation with a specific emphasis the following issues:

(A) "Developing a report for the legislature and governor, due by December 1, 2013, of best practices and recommendations regarding how teacher and principal evaluations and other appropriate elements shall inform school district human resource and personnel practices."

(B) "Taking the new teacher and principal evaluation systems to scale and the use of best practices for statewide implementation."

(C) "Providing guidance regarding the use of student growth data to assure [sic] it is used responsibly and with integrity."

(D) "Refining evaluation system management tools, professional development programs, and evaluator training programs with an emphasis on developing rater reliability."

(E) "Reviewing emerging research regarding teacher and principal evaluation systems and the development and implementation of evaluation systems in other states."

(F) "Reviewing the impact that variable demographic characteristics of students and schools have on the objectivity, reliability, validity, and availability of student growth data."

(G) "Developing recommendations regarding how teacher evaluations could inform state policies regarding the criteria for a teacher to obtain continuing contract status under RCW 28A.405.210."

(ESSB 5895, (7)(e)(i)-(v))

In addition, OSPI is also required to "submit reports detailing findings, emergent issues or trends, recommendations from the steering committee, and the pilot school districts, and other recommendations, to enhance implementation and continuous improvement of the revised evaluation systems" the legislature by July 1, 2013, and each following July 1 until the implementation transition period ends on December 1, 2016. This final, Year 3, project report is designed to serve as an information source to OSPI in developing the July 1 report for 2013.

State of the State Educator Survey

Deliverable Summary

The [Washington 2012 State of the State Educator Survey Report](#) takes stock of current educator evaluation systems across the state's school districts. Similar to AIR's 2011 State of the State Survey analysis, AIR conducted a statewide survey of Washington public school teachers, principals, superintendents, and school directors during the fall of 2012. The research questions addressed in the fall 2012 survey include the following:

1. Which aspects of the legislative requirements are districts aware of and where are the areas of misunderstanding or confusion?

2. What stage are districts at in terms of implementing and communicating about these changes?

3. How much variation exists across Washington districts in terms of implementation plans and timelines?

A total of 15,793 K–12 educators in the state (roughly one-quarter) completed the survey. The findings from the survey were presented in a March 2013 report. Subsequently, [four additional reports](#) were developed that were tailored specifically for the four educator groups that completed the survey: a [teacher report](#), a [principal report](#), a [superintendent report](#), and a [school director report](#). Additional tailored analyses of the data were provided at the request of the Washington State School Directors' Association.

Key Survey Results

The Washington 2012 State of the State Educator Survey Report, which represented 36 percent ($n = 4,406$) of teachers and 31 percent ($n = 262$) of principals, as well as district superintendents and school directors from 280 of 294 districts statewide, provided a number of useful findings, including:

- **Familiarity with the requirements of ESSB 5895 varies across educator positions. Compared with teachers, a higher percentage of district- and school-level administrators indicate that they are familiar with these legislative requirements.** Ninety-three percent of principals, 94 percent of superintendents, and 84 percent of school directors report being *Somewhat familiar* or *Very familiar* with the legislative requirements set forth by the state, compared with only two thirds of teachers.

- **The majority of districts surveyed have started a planning process around their new evaluation systems.** Nearly 73 percent ($n = 204$) of the 280 districts that responded to the survey reported that they have begun the planning process around their new teacher and principal evaluation systems.

- **Communication about teacher evaluation has been greater than communication about principal evaluation.** One hundred and five (37.5 percent) districts reported communicating information about the use of observations and student growth measures within their teacher evaluation systems compared with 32 districts (11.4 percent) that had shared information about how these measures will be used to evaluate principals.

- **Fifty percent or fewer of districts report providing districtwide training for principals and teachers around these new systems, as of fall 2012.** More than fifty percent ($n = 142$) of districts provided training on the new instructional framework for teachers, compared with 33.9 percent ($n = 95$) that provided training on the new leadership framework for principals. However, among educators who did attend district trainings, 88 percent of teachers and 90 percent of principals found them to be useful.

- **Time demands of the new evaluation system are cited as the most common concern across all educator groups.** Twenty-four percent ($n = 2,948$) of teachers, 56.3 percent ($n = 482$) of principals, 55.4 percent ($n = 195$) of superintendents, and 56.8 percent ($n = 134$) of school directors are most concerned about the potential time demands required for the new systems.

- **Educators identified (1) access to lessons learned, successes, and challenges from the TPEP pilot districts and (2) additional training around the new evaluation systems as key areas of strategic support that are needed.** Thirty-six percent ($n = 4,406$) of teachers and 30.6 percent ($n = 262$) of principals requested information about the successes and challenges of districts that have already pilot-tested new evaluation systems, although this was not cited as a specific strategy of support requested by district-level administrators. Twenty-eight percent ($n = 3,432$) of teachers, 40.8 percent ($n = 349$) of principals, 27.6 percent ($n = 97$) of superintendents, and 11.9 percent ($n = 28$) of school directors indicated that additional training to ensure rater agreement would alleviate some concerns about the new evaluation systems.

Policy-to-Practice Cohorts

Deliverable Summary

With the introduction of the new educator evaluation legislation in 2010, nine districts were awarded funding to pilot the design and implementation of new educator evaluation systems. The pilot sites' experiences and advice serve to inform OSPI's policymaking, guidance, and supports, as well as plans for the rollout of the evaluation reform statewide. The accumulated experience of the original nine TPEP pilot sites provides a critical source of expertise to inform OSPI's statewide evaluation reform implementation. Senate Bill 5895, section e (ii), recognized the critical importance of continuing to learn from district experiences by requiring that, "[i]n developing these recommendations the experiences of school districts and teachers during the evaluation transition phase must be considered." As detailed in the Executive Summary and Introduction, SB 5895 articulated several issue areas as a specific focus for ongoing recommendations.

To continue to learn from the ongoing experience of these nine pilot sites and to support ongoing efforts in the pilot districts to implement educator evaluation reforms, OSPI created three policy-to-practice cohorts. Each cohort was assigned a specific topic area from those articulated in SB 5895, which are described in Table 1.

Table 1. Policy-to-Practice Cohorts and Topics

Cohort #1: Impact on Student Learning	Cohort #2: Impact on Professional Learning	Cohort #3: Impact on Human Capital Decisions
Participating Districts: ▪ Anacortes ▪ The Consortium[2] ▪ Yakima	*Participating Districts:* ▪ Central Valley ▪ North Mason ▪ Snohomish	*Participating Districts:* 1. Kennewick 2. North Thurston *3.* Othello
Topic: the use of student growth data to ensure it is used responsibly and with integrity, particularly regarding the impact that variable demographic characteristics of students and schools have on the objectivity, reliability, validity, and availability of student growth data.	*Topic:* (1) how to take the new teacher and principal evaluation systems to scale and the use of best practices for statewide implementation; (2) refining evaluation system management tools, professional development programs, and evaluator training programs with an emphasis on developing rater reliability.	*Topic:* (1) how teacher evaluations could inform state policies regarding the criteria for a teacher to obtain continuing contract status under RCW 28A.405.210; (2) how teacher and principal evaluations and other appropriate elements shall inform school district human resource and personnel practices.

---

[2] The Consortium pilot site is comprised of nine rural school districts from the Northeast Washington Educational Service District (ESD 101). The Consortium districts that participated in the P2P meetings included Pullman School District and Medical Lake School District.

For each topic area, the policy-to-practice cohort reviewed emerging research regarding teacher and principal evaluation systems and the development and implementation of evaluation systems in other states. The cohorts met together three times as a group between September 2012 and April 2013, participated in focus group discussions, structured conversations, and conducted local "constituent discussions" in their own districts on their assigned topics. Participants also responded to an AIR survey that gathered their input on the prioritization of 38 key policy recommendations specifically for cohort 3, Impact on Human Capital Decisions.

Policy-to-Practice Meetings Overview. In Year three, OSPI convened three policy-to-practice cohort meetings with participating districts. On average, approximately 76 staff members from the nine pilot sites attended the meetings. AIR facilitated meetings of all the districts on September 19, 2012, January 24, 2013, and April 24, 2013, by providing resources and facilitation support to help districts grapple with their selected topics and to inform the state and future recommendations. AIR collected data during the meetings as pilot districts shared their insights, experiences, and perspectives on each topic area.

September 19, 2012. Participants received an overview of the national policy environment and emerging research on each of the three cohort topics. Following the overview, participants completed a jigsaw reading activity that utilized key research for their cohort topic and engaged in role-alike discussions of the research. AIR and OSPI staff next facilitated focus group discussions on each of the three cohort topics. Information from the focus group discussions are included in the key findings section for each cohort (see pp. 11-46).

In preparation for the January meeting, participating pilot districts were asked to hold local constituent discussions (similar to focus groups) with their staff to gain further insights from a broader spectrum of stakeholders. To guide these local constituent discussions, AIR provided multiple sets of discussion guide materials on each topic area, and the districts selected the format and questions that best met their needs.

January 24, 2013. The January meeting provided an opportunity for pilot districts to prepare and deliver district presentations to OSPI, AIR, and the other pilots that summarized and reflected on their constituent discussions on each topic area. Districts reviewed and discussed the draft State of the State Survey Report and a draft district communication guide (discussed on p. 47). This meeting also convened an expert panel presentation on three topic areas that intersect in important ways with educator evaluation: (1) Common Core State Standards implementation, (2) teacher certification and licensure, and (3) principal preparation. The panel featured education leaders from OSPI (Greta Bornemann), the Professional Educator Standards Board (Jennifer Wallace), the Washington Education Association (Jim Meadows), and Seattle Pacific University (Bill Prenevost).

April 24, 2013. The April meeting began with a presentation by AIR staff on human resource management that included tabletop activities to guide participants in thinking systemically about the various teacher and principal policies that contribute to effectiveness, and how these policies are connected to one another in their districts. AIR provided each participant with a 200+ page Educator Talent Management Framework, which summarized in an accessible way the research base on a range of human resource policies. Participants then broke into groups to read national and state-specific research about one of five topics related to educator evaluation and human

resource decisions: reduction in force; teacher placement (principal authority); preparation, certification, and contract status; recruitment and hiring; and mentoring and induction. Through structured discussions, each of the cohort groups then turned what they had learned from the research and from their experiences with educator evaluation reform into recommendations around their topic, including a rationale and Washington-specific considerations. AIR then conducted a survey of all 59 meeting participants to gauge the extent to which they saw these recommendations as priorities and the extent to which they believed Washington was ready to address them at present. The findings from this survey, the P2P meetings and the local constituent discussions are summarized in the next section. Meeting agendas are available at: http://tpep-wa.org/tpep-events-files/tpep-pilot-site-meetings/

Policy-to-Practice Cohort Key Findings

The sections that follow summarize the key findings and recommendations from each policy-to-practice cohort. For each cohort, a summary is provided of their topic area, the research and policy overviews they reviewed, and the combined findings from all meeting discussions, focus groups, and constituent discussions. For cohort #3, *Impact on Human Capital Decisions*, the findings are reported through an analysis of survey results describing all policy-to-practice cohort participants' perspectives on the 35 recommendations created at the April 25, 2013 meeting.

Cohort #1. Impact on Student Learning

Cohort #1 focused on two of the seven issues areas identified in SB 5895:

> (C) "Providing guidance regarding the use of student growth data to assure [sic] it is used responsibly and with integrity."

> (F) "Reviewing the impact that variable demographic characteristics of students and schools have on the objectivity, reliability, validity, and availability of student growth data."

(ESSB 5895, (7)(e)(ii))

Measuring student growth in a fair and reliable manner remains an ongoing challenge in implementing educator evaluation reforms. Moreover, ensuring that the student growth data generated is used responsibly and with integrity requires careful attention to the overall quality of the data and the strength of the evidence provided by data. Much of the research debate about using measures of student growth in educator evaluation centers on the impact of demographic characteristics of students and schools on the reliability and validity of student growth measures.

To understand these complex issues more fully, policy-to-practice cohort participants reviewed emerging research and examined information about how other states and districts are implementing measures of student growth as part of educator evaluation. This information is summarized below to provide the reader with an overview of the information that policy-to-practice cohort participants considered and reflected on as part of their conversations and recommendations.

State Policy and Practice Overview. States have adopted a wide range of approaches to measuring student growth for use in educator evaluation. This diversity reflects the range of teaching assignments, subject areas, and assessment types that currently exist in most states and districts. Although much of the debate over student growth in educator evaluation centers on the use of state standardized assessments, researchers estimate that a full 69 percent of teachers provide instruction in subjects not currently assessed through standardized tests (Prince, Shuermann, Guthrie, Witham, Milanowski, & Thorn, 2009).

As Table 2 demonstrates, most states require or recommend a combination approach that utilizes a value-added or percentile model in grades and subject areas that have state-wide standardized assessments, and a student learning objective[3] approach—either in all subject areas and grades, or in the subject areas and grades not assessed through standardized tests. Other states rely entirely on value-added approach, which poses significant measurement challenges and often requires developing new standardized assessments to cover a wider range of courses and grades.

Table 2. State Approaches to Student Growth Models

Approach	Number of States
1. State-determined value-added or percentile model (including non-tested subjects and grades).	3
2. Student learning objectives or similar growth goal-setting approach (all educators).	4
3. Combination of value-added or percentile models and student learning objectives.	19

Source: Center on Great Teachers and Leaders, 2013

Emerging Research Summary. Currently, no consensus exists within the research community about the use of student growth measures in educator evaluation. Strong divides exist among researchers on a range of issues, such as (1) the appropriateness of utilizing state standardized assessments and advanced statistical models as part of high-stakes performance evaluation and (2) the rigor and comparability of student learning objectives (SLOs) as well as the time burden on teachers who must implement them (Kane and Darling-Hammond, 2012; Darling-Hammond 2012; Roth, Huffcutt, & Bobko , 2003; Alexander, Entwisle, & Olson, 2007; Newton, Darling-Hammond, Haertel, & Thomas, 2010; McCaffrey, Koretz, Lockwood, & Hamilton, 2005; Goldhaber & Theobald, 2012; Jackson, 2012; Harris & Anderson, 2012). Although a full discussion of these research debates is beyond the scope of this report, there are several considerations that researchers largely agree on as critical for states and districts to take into account when measuring student growth.

- **All measurement, regardless of the model used, is subject to bias and error**. There are no "perfect" measures of student growth, and all measurement models provide only an approximation of a student's true learning and growth in the classroom, and a very partial picture of educator performance (Raudenbusch and Jean, 2012; Kane and Darling-Hammond, 2012). For this reason, utilizing any student growth measure for high-stakes

---

[3] In some states, this type of measure is referred to as a "student growth objectives" (e.g., Colorado) or "student learning targets" (e.g., Louisiana) measure.

personnel decisions should be done with great caution and only in conjunction with evidence from other measures of educator performance.

- **Using multiple measures of student growth from multiple points in time can provide an important check on the consistency and strength of the evidence under consideration.** Existing research has focused largely on correlations between value-added or percentile models and measures of teacher practice, such as classroom observation scores or student surveys (MET Project, 2013). New research, however, is slowly emerging that examines the degree of consistency between *multiple measures of student growth.* Early findings suggest that different approaches to measuring student growth, when used in tandem, can provide a more complete and meaningful source of evidence for considering teacher impact. In their multi-year studies of performance compensation reform in Denver, Colorado and in Charlotte-Mecklenburg, North Carolina, researchers at the Community Training and Assistance Center (CTAC) found that *high-quality* SLOs were consistently correlated with higher student achievement in mathematics and reading (CTAC 2004; CTAC 2013). Utilizing multiple measures of student growth, including multiple models, assessments, and growth targets, during several years provides a check on the strength of the evidence under consideration. This helps reduce the likelihood that an individual educator will be unfairly evaluated using invalid or unreliable evidence.

- **All measurement models are only as good as the assessments on which they rely.** Measurement models are a method for analyzing assessment data and comparing data from different points in time. If the underlying assessment is poorly designed or not aligned to classroom content, the measurement approach will produce similarly poor or biased results (Lachlan-Haché, Cushing, & Bivona , 2012). Supporting educators by providing access to high-quality assessments, as well as improving assessment literacy among educators are crucial foundations to measuring student growth. Improving assessment literacy ensures district and school leaders can identify high-quality, rigorous assessments for use and can work with staff in selecting or developing assessments for both formative and summative purposes (Kane & Darling-Hammond, 2012).

- **For measures of student growth to contribute to professional growth and increased student learning—standardized assessment data alone is an insufficient source of information.** Standardized assessments provide teachers and principals with information about their students' overall progress on meeting content standards each year. The limited frequency of, and lengthy delay in, receiving test results, however, make this data insufficient for informing immediate changes to school leadership and instructional practice. Educators need access to a range of formative and summative assessments to engage in ongoing data analysis about the progress of the students currently in front of them (Goe, 2008; Goe, Bell, & Littlel, 2008).

Policy-to-Practice Key Findings: Cohort #1

In September 2012 and January 2013, pilot districts participating in the policy-to-practice cohort meetings reviewed the existing research on measuring student growth, and shared their progress on incorporating student growth into their teacher evaluation system as well as any advice they had for OSPI and for other districts planning for implementation. As part of this discussion, district

leaders reported on what they heard in their constituent discussions and shared hopes, concerns, and areas where support was still needed in order to measure student growth in a fair and reliable manner.

- **Implementation Progress**
  - **Cautions.** Many districts were still in the process of selecting an approach to measuring student growth. Participants from these districts noted that they had been wary of starting too far down one path out of concern that, as had happened with Senate Bill 5895, the legislature would revise the requirements and mandates around measuring student growth again.
  - **Emerging efforts.** At the September and January meetings, a few pilot districts reported progress on using student growth goal-setting as part of the teacher evaluation process. Principals were working with teachers to identify appropriate assessment data and to craft rigorous, appropriate growth targets for their students. Several districts had begun to review OSPI's student growth rubrics and to begin planning for how to incorporate them into the districts' summative scoring process.

- **Concerns.** Districts raised a range of concerns about using student growth in teacher and principal evaluation.
  - **State testing calendar.** Districts expressed frustration that state assessment data is typically not released until well after the current school year has ended, which does not fit well with educator evaluation timelines that require a summative evaluation score in the spring. This lack of alignment in the timing also makes the data difficult to use to improve teaching and leadership practice through timely feedback.
  - **Clarity of purpose.** Districts were concerned that teachers will think that student growth can be measured only through "tests" and will miss the opportunity to measure student growth through a wider range of assessment approaches that may better fit their classrooms or courses, such as performance-based assessments.
  - **Educator misperceptions.** Districts were concerned that focusing too heavily on standardized assessments would result in teachers viewing the new evaluation system as just one more policy reform aimed at limiting instructional time in favor of more testing.
  - **Fairness and consistency.** Districts were unsure how OSPI would ensure fairness and consistency across schools and districts in the use of the student growth rubrics and setting growth targets.

- **Supports Needed.** Districts identified several concrete areas where crucial supports are needed in order to implement student growth measures effectively.
  - **Communication resources:** the shift from a focus on student achievement (or attainment) to student growth (change over time). Districts need practical tools and methods for communicating this shift to educators, parents, and the wider community.
  - **Assessment planning:** guidance and direction on assessment planning in light of Common Core State Standards. How can districts be proactive in ensuring that the next generation of assessments aligned to the Common Core State Standards will be more directly useful for adjusting and improving instructional practice?

- **Business rules development**: creating business rules for student growth. Districts need guidance on how best to address the problem of student mobility, particularly in classrooms where turnover throughout the duration of the course involves the majority of students.

- **Assessment literacy**: access to time and resources for improving teachers', principals', and district administrators' level of assessment literacy. Because the majority of teachers provide instruction in grades and subject areas for which no state standardized assessments exist, knowing how to identify and create rigorous, aligned assessments in a range of styles (e.g., performance-based) is a core lynchpin and pre-requisite for generating better student growth data.

- **Student growth goals:** setting rigorous, but attainable, student growth goals that are fair and consistent across schools and districts. Both teachers and leaders need extensive training and support to do this well. A statewide bank of vetted student growth goals across a range of subject areas and grades that can serve as exemplars for districts is a vitally needed resource.

- **Data management systems:** supporting cross-district collaboration to identify management systems for student learning data. In the absence of a statewide data system, districts need to leverage each other's expertise and resources in order to manage assessment data.

- **Advice to Other Districts.** Policy-to-practice cohort participants discussed several early lessons their teams had discovered as they implemented student growth measures.

  - **Differentiate flexibility**. The degree of flexibility given to teachers around student growth goals might need to be differentiated based on the teacher's performance level and experience. Less experienced teachers or teachers who are struggling may require more fixed options and direct guidance for setting growth targets, whereas experienced, effective teachers can be granted greater flexibility.

  - **Leverage teacher leaders.** Principals are already overwhelmed just implementing the teacher practice elements of the new evaluation systems. Leveraging teacher leaders to support principals in assisting other teachers with student growth goal-setting, and selecting or creating assessments, is one option for preventing principal overload and burnout.

  - **Cultivate alignment.** Create a clear alignment between teachers' analysis of their students' needs, their student growth goals, and their focus in the instructional framework. If these components are aligned, teachers will have a more cohesive, focused evaluation experience.

  - **Rethink and refocus.** Encourage teachers and principals to think of the student growth goal-setting process as an opportunity to better focus on the students who need the most help. Student growth goals should address the needs of students who face the most significant barriers in terms of poverty, lack of support in the home, discipline or attendance issues, or language barriers.

- **Advice to OSPI, the TPEP Steering Committee, and the State Legislature.** The overwhelming advice that policy-to-practice cohort participants raised for state education

leaders was to continue to prioritize district and educator flexibility in designing approaches to measuring student growth.

- **Prioritize teacher empowerment/autonomy**. Keep teacher empowerment and autonomy in the student growth measurement process at the forefront of all recommendations and policymaking on student growth. Flexibility empowers teachers to identify and focus a set of goals on the students in their classrooms who need the most help. Flexibility also encourages teachers to seek out support and expertise from their colleagues, which will support collaboration among peers.

- **Prioritize district flexibility.** Continue to give districts the flexibility to select the assessments and measurement approaches that make sense for their schools and staff.

- **Hopes.** Pilot districts expressed a number of hopes about the possible positive outcomes that could emerge from measuring student growth as part of educator evaluation.

  - **Better assessments.** Districts hoped that the requirement to measure student growth will encourage the development and use of better assessments, particularly more formative assessments that will better inform instruction.

  - **Better data use.** Districts hoped that teachers and leaders alike will be able to learn to use data more effectively rather than simply having more data to sort through.

  - **Embedding ongoing data use in educator practice.** Several districts expressed a hope that the student growth goal-setting process will help to embed ongoing, consistent data use within individual teacher practice, teacher teams' practices, and school and district leadership team practices.

Cohort #2. Impact on Professional Learning

Cohort #2 focused on two of the seven issues areas identified in SB 5895, specifically:

> (B) "Taking the new teacher and principal evaluation systems to scale and the use of best practices for statewide implementation."

> (D) "Refining evaluation system management tools, professional development programs, and evaluator training programs with an emphasis on developing rater reliability."

(ESSB 5895, (7)(e)(ii))

The impact of educator evaluation on professional learning is a large topic area that includes initial professional development focused on evaluation system training all the way to aligning professional development identification and delivery systems with evaluation system results. For Year 3, policy-to-practice cohort #2 was given a more targeted task of considering approaches to professional learning for evaluation system training, especially those that can best promote strong rater reliability and agreement.

Reflecting on their experiences designing and implementing new evaluation systems since 2010, Washington's pilot participants described rater agreement as a "huge" issue that is intertwined with educators' trust in the system. The statewide survey corroborated this finding; Washington State teachers' greatest worry about the new evaluation was that their performance might not accurately be rated, with more than 40 percent of teachers sharing this concern (Brown-Sims,

Clayton, Chen, & Brandt, 2013). Although the key findings from the cohort extend beyond narrow recommendations on rater reliability and agreement alone, the research and policy overviews the participants considered were focused closely on evaluator training and rater agreement. The information below summarizes the emerging research as well as current state policies and practices around evaluator training and rater agreement. This information is included to provide the reader with appropriate context for understanding the conversations and recommendations that were made by the cohort participants.

Emerging Research Summary. High levels of rater agreement and reliability are critical for ensuring a fair, credible, useful, and legally defensible teacher and principal evaluation system. Evaluators unable to achieve high levels of rater agreement create an unfair system that is not particularly useful for making important personnel decisions or for informing professional development planning. Low levels of reliability and agreement also open evaluation systems up to legal challenges by teachers and principals who are being dismissed on the basis of incompetence, particularly if they feel the system is unfair or invalid (Pullin, 2011). Trust in the system is critical; thus districts should do everything they can to ensure high levels of rater reliability and agreement. Time spent upfront to maximize rater agreement is likely to save districts significantly more time down the road defending ratings.

Achieving high levels of rater agreement depends partly upon using well-designed rubrics for observation and rating (Melchers, Lienhardt, VonAarburg & Kleinmann, 2011; Graham, Milanowski, & Miller, 2012; Lumley, 2002); however, it depends most heavily on the quality and frequency of evaluator training and calibration. High-quality evaluator training can take up to a full week of intense orientation to and discussion about the evaluation standards, rubric components, tools, and processes; practice rating using multiple videos or live observations; feedback and discussion on appropriate ratings; and assessment of rater skill levels, which can also take several hours. After evaluators undergo this initial training, they should continue to receive one to two days of ongoing calibration training every year to protect against rater drift. Training can take place in person, online, or as a combination of the two. Online training can be less costly and, therefore, more frequent; yet face-to-face training offers greater opportunity for collaboration and discussion about evaluation ratings and processes (Daly, 2011; Fry & Ramsdell, 2011; McClellan, Atkinson, & Danielson, 2012; Hamman, Beaubien, & Holt, 1999; Conway, Jako, & Goodman, 1995).

The Measures of Effective Teaching (MET) Project, funded by the Bill and Melinda Gates Foundation, attempted to test measures of effective teaching by working with 3,000 volunteer teachers in seven school districts over three years. The study examined classroom observation instruments, student standardized test scores using a value-added model, and student surveys. All volunteer teachers taught in elementary or middle subjects and grades and the evaluation data gathered had no consequences for the teachers who volunteered. Although a full discussion of the scope, limitations, and findings of this study are beyond the purposes of this report, as a massive research undertaking, the MET Project emphasized the importance of training all classroom observers and discovered a number of interesting findings in the process. These findings can help inform state and district decision-making around the use of observers, the number and length of observations, and the types of training needed. Specifically, the research team found the following:

- Having two observers instead of one increases reliability significantly more than having the same observer score more lessons.

- A potential cost-effective way of increasing reliability is to have additional observers watch at least part of a lesson; observations based on the first 15 minutes of lessons were about 60 percent as reliable as full-lesson observations.

- It is important to have at least two full-length observations because many indicators on the rubric will not be evident in the first 15 minutes.

- School administrators tended to rate teachers higher than did external observers (from outside the school).

- School administrators discerned greater differences in teaching practice than did external observers.

- Having observations conducted by external observers can serve as a check against in-school bias.

The study notes that relatively high reliabilities can be achieved in various ways, which means that school districts will need to decide how to allocate time and resources to classroom observations. The following observation configurations provide some illustrations:

- The reliability of a single 45-minute lesson observed by the teacher's own administrator was lowest (.51) of the seven observation configurations tested in the MET study.

- Two 45-minute lessons observed by different administrators had a reliability of .67; one lesson observed by the teacher's administrator combined with three 15-minute lessons observed by three additional peer observers produced the same reliability.

- Two 45-minute lessons observed by the teacher's own administrator, one lesson observed by a peer observer, and three 15-minute lessons observed by three additional peer observers produced the highest reliability (.72) of the seven configurations.

State Policy and Practice Overview. Although very few states have legislated requirements for rater reliability and agreement, a number of states do hold districts responsible for training their evaluators for this purpose. A review of state policy reveals the following:

- Iowa: evaluators must take an online renewal course every five years and complete two online professional development courses that align to district goals.

- Kentucky: evaluators must complete 12 hours of training every two years.

- North Carolina: experienced evaluators must continue to attend a half-day in-person workshop at the start of each school year on rater agreement and implementation fidelity.

- Washington, DC: all evaluators must be retrained each summer and participate in calibration exercises throughout the year.

To ensure rater agreement, and protect against rater drift, experts recommend *annual* testing of evaluators as best practice even among the most expert and experienced evaluators (McClellan et al., 2012). Such evaluator proficiency tests are common practice and in some districts and states are embedded in a system of evaluator certification and re-certification. For example, evaluators in Delaware, Tennessee and Houston must pass a proficiency exam prior to conducting

evaluations. Pittsburgh Public Schools' use of the Instructional Quality Assurance Certification Process includes a two-level evaluator training structure. At Level 1, the training process requires evaluators to assess five 30- to 60-minute videos and demonstrate established levels of rater agreement, evidence of objectivity and alignment to rubric components, and accuracy of assessment. At Level 2, evaluators must engage teachers in inquiry-based conversations and offer actionable next steps informed by the evaluation results. Evaluators must pass a written test of curricular knowledge, observe three classrooms with a master educator, and provide a written postobservation debrief to obtain certification (Curtis & Weiner, 2012).

Policy-to-Practice Cohort Key Findings: Cohort #2

In September 2012 and January 2013, pilot districts participating in the policy-to-practice cohort meetings reviewed the existing research on evaluator training and shared any advice they had for OSPI and for other districts planning for evaluation training. As part of this discussion, district leaders reported on what they heard in their constituent discussions and shared ideas for what supports districts needed from the state in order to ensure evaluation training could be meaningfully carried out for both evaluators and the educators being evaluated.

Supports Needed

- **Time and resources.** To implement evaluation training through an ongoing, meaningful process that fits school needs, districts need access to more time and sustained resources. Specifically, a financial block grant would enable planning for customized, district selected professional development to be long term. Time and resources are practical prerequisites for evaluation system implementation and scaling up.

- **Ongoing professional development.** Educators need opportunities for ongoing professional development, including evaluator training. It is critical that this training be more extensive than a single day or single event and be embedded in existing collaborative structures within the school.

- **Individualized, job-embedded professional development options.** Districts need access to professional development options that can be individualized by teacher and principal needs, and goals that are identified as part of the evaluation process. Teachers and principals need professional development that can be accessed quickly and implemented through job-embedded approaches that allow for a full cycle of learning new information, implementing changes in practice, and reflecting on the change, and adjusting practice in response to reflections.

- **Training on coaching.** Principals and administrators need professional development and training on how to coach the teachers and principals they supervise and how to have constructive conversations about instructional and leadership practice.

Advice to Other Districts

- **Communicate with ALL your staff about evaluator training.** Extensive evaluator training has the potential to alleviate educator anxiety about evaluation reform. One pilot district noted that constituent discussions with their teachers revealed that the anxiety

about the new educator evaluation system was far greater than anticipated, and at the heart of this was concern about equity in observations across school buildings. Rigorous, high-quality evaluator training is often seen as the only approach to address this widespread concern. Make sure you communicate to your staff how you have trained evaluators and how they have demonstrated their reliability and skill.

- **Train principals and teachers together.** Providing at least some common training time together opens up opportunities to hear each other's perspectives and ensure consistency in the training and messages received across both groups.

- **Evaluators want training.** Evaluators will likely embrace rather than feel burdened by training to keep their judgment in check. Although training represents an additional responsibility in educators' already full schedules, it is an investment of time that will likely be seen as a priority.

- **Talk to your school board.** School boards often are equally concerned about the need for high-quality evaluator training. Therefore, school districts should not be shy about approaching their school boards for support with this aspect of educator evaluation reform.

- **Build in collaborative calibration activities.** Evaluators benefit greatly from opportunities to collaborate and practice alongside other evaluators, discuss and compare ratings, and improve their calibration. These calibration opportunities can be integrated fairly easily into regular existing principal meetings, such as at principal professional learning communities or during professional development days.

Advice to OSPI and the State Legislature

- **Resources and flexibility.** Prioritize giving districts resources and the flexibility to use resources to identify the evaluation professional development best suited to their district's unique needs. Provide monetary resources directly to districts so that district leaders can identify professional development for evaluation training that best suits the needs of their district.

- **Support within-school collaboration.** Provide support for continued educator collaborations, such as funding to provide classroom coverage for teachers to participate in Learning Walks and to visit each other's classrooms, so educators can observe teacher practice and discuss what was observed against the instructional rubric.

- **Redistribute or remove principal administrative tasks.** Redistribute or remove less critical reporting and administrative tasks from principals' workloads to provide the time and human capital resources necessary to fully implement the educator evaluation system with fidelity.

- **Reinstate learning in-service days**. Dedicated time for teachers to dig more deeply into the instructional framework as teams of educators is crucial for improving teacher practice and encouraging teacher buy-in to support new evaluation systems. Learning in-service days are one resource that would be an enormous support for districts.

Cohort #3. Impact on Human Resource Decisions

Cohort #3 focused on two of the seven issues areas identified in SB 5895, specifically:

(A) "Developing a report for the legislature and governor, due by December 1, 2013, of best practices and recommendations regarding how teacher and principal evaluations and other appropriate elements shall inform school district human resource and personnel practices."

(G) "Developing recommendations regarding how teacher evaluations could inform state policies regarding the criteria for a teacher to obtain continuing contract status under RCW 28A.405.210."

(ESSB 5895, (7)(e)(ii))

At the April 24, 2013, policy-to-practice meeting, AIR provided each participant with a 200+ page Educator Talent Management Framework, which outlines the large national research base on human capital management, highlighting in particular the interconnectedness among various human resource policies. The Framework presents the ways in which teacher and principal evaluation reform is part of a larger system of teacher and principal effectiveness that includes links between evaluations and other aspects of the educator's career continuum including: preparation, recruitment, hiring, induction and mentoring, professional development, compensation and incentives, and school environment. The Educator Talent Management Framework provides summaries of existing research on a broad range of human capital management topics. Readers should note, however, that rigorous research examining linkages between the use of evaluation results in human resource decisions (e.g., continuing contract status, reductions in force, and staffing assignment) and impacts on student learning are very limited.

The range of additional human resource functions that are tied to teacher and principal evaluations are many, including continuing contract status, reductions in force, dismissal, promotion, career ladders, equitable teacher distribution, and other policies and practices. States and districts across the country have only very recently begun to implement new evaluation systems and only a few districts have engaged in reforms linking evaluation results to human resource decisions. Those of particular relevance in the Washington context were explored most deeply through literature reviews, discussions among the pilot steering committees and the wider body of staff in their districts, focus groups, and surveys. During the meeting, policy-to-practice cohort participants developed a set of 35 recommendations on human resource decisions and educator evaluation.

The following sections provide a summary of the research and national policy that cohort participants reviewed on continuing contract status, reductions in force, staffing assignments, and additional human resource decisions. This information is included to provide the reader with appropriate context for understanding the conversations and recommendations among the cohort participants. Next, the report presents the results of a post-meeting survey that asked participants to identify which of the 35 recommendations created at the meeting were of a high priority and which they felt the state was most ready to address.

Continuing Contract Status

State Policy and Practice Overview. The purpose of continuing contract status for teachers is twofold: (1) to provide an element of job security that will make the profession more attractive,

thereby improving teacher recruitment and retention; and (2) to provide an element of job security that will protect teachers from political forces that could cost them their jobs — if, for example, a controversial subject is taught in a way that diverges from the views of the principal, superintendent, or community (Hassel, Kowal, Ableidinger, & Hassel, 2011).

Many states have passed educator effectiveness legislation that includes new requirements around continuing contract status. As of September 2012, most states, similar to Washington, require a minimum of three years of teaching experience before continuing contract status is awarded. As illustrated in Figure 1, five states require two years of experience, five states require four years of experience, and six states require five years of experience. Whereas in many cases recent legislation increased the number of years for tenure eligibility, in Ohio the length of service requirement was reduced from seven to five years. At the extremes, Mississippi requires just one year of experience, whereas Florida and Rhode Island allow only annual contracts for teachers (NCTQ, 2012; NCTQ, 2010a).

Figure 1. State Requirements for Teachers to Obtain Continuing Contract Status
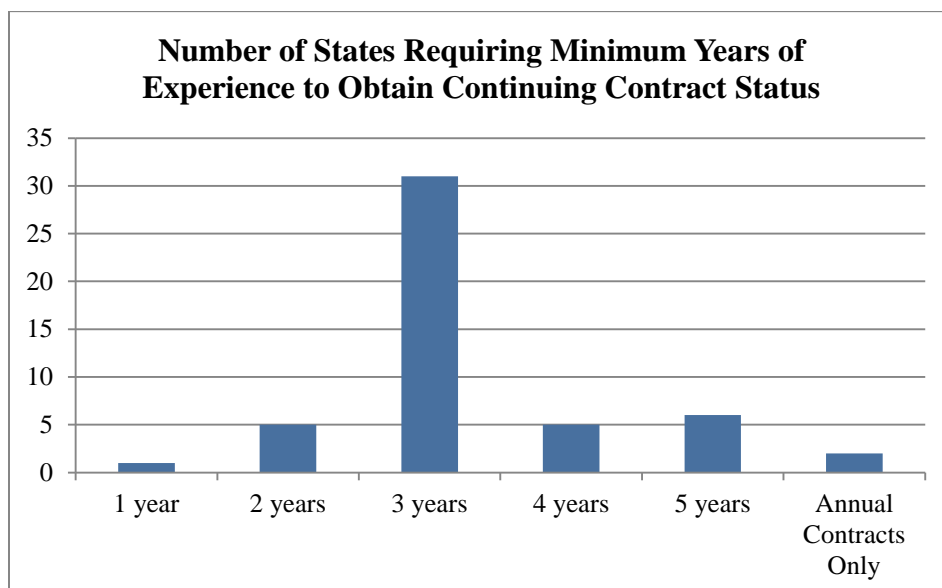
Prior to the recent wave of legislation across the country, contract status was based largely on years of teaching experience, although teaching effectiveness may factor into the decision to grant continuing contract status depending upon the district. Continuing contract status was, therefore, seen as easily obtained, if not automatic. Teachers who have attained continuing contract status tend not to be dismissed. For example, TNTP reported that 86 percent of school administrators do not pursue the dismissal of ineffective teachers even when they believe it is warranted, citing the time-consuming process of so doing (Weisberg, Sexton, Mulhern, and Keeling, 2009). Meanwhile, the Center for American Progress notes that, in the rare cases that tenured teachers are dismissed, it is for egregious conduct as opposed to poor performance (McGuinn, 2010).

In recent years, nine states (Colorado, Delaware, Florida, Louisiana, Michigan, Nevada, Oklahoma, Rhode Island, and Tennessee) have created requirements that continuing contract decisions be based predominantly on evidence of student learning and seven additional states

(Connecticut, Hawaii, Illinois, Indiana, Massachusetts, New Jersey, and New York) have required that evidence of student learning be considered when determining contract status (NCTQ, 2012). Meanwhile, large school districts such as Washington, DC and New York City have dismissed or encouraged the departure of large numbers of teachers (TNTP, 2012).

Emerging Research Summary. Currently, the volume and rigor of research examining the impact of these new policies for continuing contract status is insufficient to inform state decision making (McGuinn, 2010). One trend has been noted: from 2008–09 to 2009–10 alone, the percentage of teachers who left the profession because their contracts were not renewed, rather than for other reasons, increased from 27 percent to 35 percent. The percentage of teachers who moved schools because their contracts were not renewed similarly increased from 21 percent to 31 percent (U.S. Department of Education, 2011). No studies, however, of how these changes influenced student outcomes are currently available.

Reductions in Force

Compared to continuing contract status, at the national level even fewer policy changes have been made that specifically addressed Reductions in Force (RIF). Moreover, no research currently exists on the impact of RIF policy changes on student outcomes.

State Policy and Practice Overview. Fourteen states require that seniority be the primary factor in determining teacher layoffs, whereas the large majority of states leave this decision to local school districts (NCTQ, 2010b). Looking at the district level, all 75 of the nation's largest school districts factor in seniority to determine the order of layoffs, and in more than 70 percent of these districts, seniority is the sole determinant (Goldhaber, 2011). The state of Arizona is unique in its prohibition of the use of seniority in the determination of RIFs. Missouri requires that seniority be the primary determinant for provisional contract teachers, but that for continuing contract teachers both seniority and performance be considered when layoffs must be made (NCTQ, 2010b).

Staffing Assignments

Staffing assignments have relevance both for transfer eligibility and for the equitable distribution of teachers for students from minority or high-poverty backgrounds. As noted below, most policy development in this area has taken place at the district-level. In general, the research base on one aspect of staffing assignments—the equitable distribution of effective educators— provides some guiding information for consideration.

State Policy and Practice Overview. Regarding transfer eligibility, little policy development has taken place at the state level, although some districts have enacted reforms in recent years. Congruent with Washington's own current law, teachers in Houston Independent School District with unsatisfactory performance are prohibited from transferring, thereby precluding underperforming teachers from transferring across schools without ever shoring up their deficiencies. In contrast, Hillsborough County Schools in Tampa, Florida implemented an approach that gives discretion to the principals at potential receiving schools (Lemke, Thomsen, Wayne, & Birman 2012). Principals in search of a transferring teacher receive a profile of the

prospective transfer. The profile summarizes the teacher's value-added score and teaching history. The principal can then decide whether to accept the transfer. No research on the impact of the policy described in these two cases is currently available.

Emerging Research Summary. Some researchers have cautioned against forced placements that undermine principals' ability to build a cohesive instructional staff, and, indeed, several districts have eliminated forced placement, but this body of research is limited (Franck, Kellihers, &Varghese, 2011).

Regarding equitable distribution, or the placement of high-performing teachers in low-performing schools, relevant research does exist to inform state policymaking. Research finds that, contrary to popular opinion, transferring effective teachers from a low-need setting to a higher-need setting does not on the whole impact their ability to be effective. In their analysis of seven to 11 years of teacher performance data from North Carolina and Florida, researchers from the National Center for Analysis of Longitudinal Data in Education Research found that high-performing teachers who moved to low-performing schools either maintained or improved performance during the years after the move. In fact, movement between advantaged and less advantaged schools—in either direction—was associated either with the maintenance of teacher performance prior to the move or slight improvement in performance (Zeyu, Ozek, & Corritone, 2012).

Equitable distribution policies may take the form of financial incentives, forced transfer, or the reassignment of cohorts of teachers and, in some cases, a principal. One notable example of equitable distribution worth highlighting is Charlotte-Mecklenburg Schools' Strategic Staffing Initiative (SSI). Initiated in 2008, under the SSI, the lowest-performing, most challenged schools became the first in line to receive teachers and principals with demonstrated track records of success. High-performing principals selected a team (consisting of an assistant principal, a literacy facilitator, a behavior management technician, and up to five teachers with proven success) with whom to transfer to a high-need school. To incentivize teachers, the district offered $10,000 recruitment bonuses and retention bonuses worth $5,000 in the second and the third years (Charlotte-Mecklenberg Schools, 2009). The first year results were favorable: SSI schools (i.e., those that received high-performing teachers) showed increases in student proficiency on assessments in reading, math, and science (Travers & Christiansen, 2010). By 2011, about half of the seven participating schools outpaced comparison schools in student growth, with the greatest increases in student growth at schools with where the new team represented a larger portion of total faculty members, and schools with high teacher turnover rates (Schoeneberger & Pulliam, 2011).

The Talent Transfer Initiative (TTI), a randomized control trial study funded by the U.S. Department of Education Institute of Education Sciences, provides bonuses to the highest-performing teachers (measured in terms of value-added student growth scores) in certain categories (e.g., math) of $20,000 to transfer and remain for at least two years in a selected low-achieving school in their district. At the time of writing, the impact of this initiative over time on student achievement was still being evaluated and will be released in the final study report; in this interim report, however, researchers did find that the incentives were sufficient to fill 90 percent of the vacancies in the seven participating districts' high-need schools with effective teachers (Glazerman, Protik, Bruch, & Seftor, 2012).

Policy-to-Practice Key Findings: Cohort #3

As noted above, at the September and January policy-to-practice meetings, pilot districts reviewed research, participated in a focus group discussion, and shared the results of their constituent discussions. At the April policy-to-practice meeting, participants reviewed research reports and specific Washington documents on five priority human resource topic areas related to educator evaluation. Participants broke into groups to develop recommendations on one of five priority topics: reduction in force; teacher placement (principal authority); preparation, certification, and contract status; recruitment and hiring; and mentoring and induction. Thirty-five recommendations emerged from group based on the accumulated discussions across the three meetings. AIR then conducted a survey to determine the extent to which participants viewed all 35 recommendations as: (a) a priority, and (b) an area where the state was ready to enact those recommendations. The information presented below summarizes the survey results.

Survey Methodology. Between May 17, 2013 and May 27, 2013, AIR invited 60 members of the WA TPEP Policy-to-Practice group to complete a 25-minute online survey about their perspectives on the recommendations that emerged from participant activities during the meeting. The 11-item survey consisted of both closed and open-ended questions. One reminder notice was sent during the survey administration window. A total of 43 respondents (71.7 percent) completed the survey. Results from the closed survey questions are reported below.

Preparation, Certification, and Contract Status. The policy-to-practice cohorts developed eight recommendations on this topic area. Respondents were asked to rate the priority of each recommendation. Thirty-nine survey respondents, or 90.6 percent, answered this question.

Figure 2 presents each recommendation in order of priority from highest to lowest. The top three recommendations that policy-to-practice cohort participants identified as "very much a priority" each focused on strengthening coherence and alignment across preparation, certification, and contract status.

- The highest priority recommendation (R1) advises state policymakers to focus on revising and simplifying the teacher licensure rules rather than expanding them: "*Do not make any additional licensure rules; any change should be limited to fixing and aspects of the existing system that are unclear and/or simplifying the current system*" (82.1 percent[4] or 32 respondents).

- The second highest priority recommendation (R2) advises state policymakers to look across systems to *"[a]lign and integrate systems for preparation, certification, and contract status to eliminate redundancies and ensure coherence"* (76.9 percent or 30 respondents).

- The third highest priority recommendation (R3) advises state policymakers to *"[r]educe the number of times new teachers must prove initial competence (e.g., teacher*

---

[4] All percentages reported for survey item response categories (e.g., "very much a priority") are calculated by dividing the total number of respondents who selected the category by the total number of respondents who answered the question. For example, in this first question, 39 people responded to the question and 32 of those 39 respondents (or 82.1 percent) identified R1 as "very much a priority."

preparation program, testing (West-E), pro- teach, district/annual evaluation” (69.2 percent or 27 respondents).

- The fourth highest priority recommendation (R4) advises state policymakers to *“[d]elay any action around preparation, certification, and contract status until there is sufficient funding to support it”* (66.7 percent or 26 respondents).

Figure 2. Prioritization of Recommendations on Preparation, Certification, and Contract Status

Recommendation	Not a priority	Somewhat of a priority	Very much a priority
R1. Do not make any additional licensure rules; any change should be limited to fixing and aspects of the existing system that are unclear and/or simplifying the current system	2.6%	15.4%	82.1%
R2: Align and integrate systems for preparation, certification, and contract status to eliminate redundancies and ensure coherence	7.7%	15.4%	76.9%
R3. Reduce the number of times new teachers must prove initial competence (e.g., teacher preparation program, testing (West-E), pro- teach, district/annual evaluation)	12.8%	17.9%	69.2%
R4. Delay any action around preparation, certification, and contract status until there is sufficient funding to support it	12.8%	20.5%	66.7%
R5. Provide immediate block grant funding to districts to determine how best to use in their context	17.9%	35.9%	46.2%
R6. Provide support for student teaching	23.1%	43.6%	33.3%
R7. Create reciprocal certification with other states	20.5%	53.8%	25.6%
R8. Do not link evaluation ratings and continuing contract status	41.0%	43.6%	15.4%

Legend: ■ Not a priority ■ Somewhat of a priority ■ Very much a priority

Next, participants were asked to rate the readiness of the state to address each of the eight recommendations. Thirty-seven survey respondents, or 86 percent, answered this question. Figure 3 summarizes the participants' responses. The top four recommendations that policy-to-practice cohort participants said the state was "very ready to address" were the same recommendations participants identified as very high priority, although their rankings are in a slightly different order.

- The recommendation to focus on revising and simplifying the teacher licensure rules rather than expanding them (R1) was both the highest priority and also the recommendation that participants most thought the state was most ready to address (75.7 percent or 28 respondents).

- The recommendation to reduce the number of times a teacher must prove initial competence (R3) was rated as the third highest priority and the second recommendation the state was most ready to address (64.9 percent or 24 respondents).

- The recommendation to delay action on certification, preparation, and contract status until support funds were available (R4) was rated as the fourth highest priority and the third recommendation the state was most ready to address (59.5percent or 22 respondents).

- The recommendation to align preparation, certification, and contract status requirements (R2) was rated as the second highest priority but ranked fourth as the recommendation the state was most ready to address (48.6 percent or 18 respondents).

Figure 3. Readiness to Address Recommendations on Preparation, Certification, and Contract Status

R1. Do not make any additional licensure rules; any change should be limited to fixing and aspects of the existing system that are unclear and/or simplifying the current system — Not yet ready to address: 2.7%, Somewhat ready to address: 21.6%, Very ready to address: 75.7%

R3. Reduce the number of times new teachers must prove initial competence (e.g., teacher preparation program, testing (West-E), pro- teach, district/annual evaluation) — Not yet ready to address: 5.4%, Somewhat ready to address: 29.7%, Very ready to address: 64.9%

R4. Delay any action around preparation, certification, and contract status until there is sufficient funding to support it — Not yet ready to address: 10.8%, Somewhat ready to address: 29.7%, Very ready to address: 59.5%

R2. Align and integrate systems for preparation, certification, and contract status to eliminate redundancies and ensure coherence — Not yet ready to address: 10.8%, Somewhat ready to address: 40.5%, Very ready to address: 48.6%

R5. Provide immediate block grant funding to districts to determine how best to use in their context — Not yet ready to address: 16.2%, Somewhat ready to address: 51.4%, Very ready to address: 32.4%

R8. Do not link evaluation ratings and continuing contract status — Not yet ready to address: 29.7%, Somewhat ready to address: 40.5%, Very ready to address: 29.7%

R6. Provide support for student teaching — Not yet ready to address: 24.3%, Somewhat ready to address: 51.4%, Very ready to address: 24.3%

R7. Create reciprocal certification with other states — Not yet ready to address: 32.4%, Somewhat ready to address: 45.9%, Very ready to address: 21.6%

■ Not yet ready to address　■ Somewhat ready to address　■ Very ready to address

Recruitment and Hiring. The policy-to-practice cohorts developed 10 recommendations on this topic area. Respondents were asked to rate the priority of each recommendation. Thirty-six survey respondents, or 83.7 percent, answered this question. Figure 4 presents each recommendation in order of priority from highest to lowest. The top four recommendations that policy-to-practice cohort participants identified as a "very much a priority" represent three fairly distinct policy areas connected with recruitment and hiring: teacher preparation, funding and budgeting, and teacher endorsements.

- The highest priority recommendation (R1) advises state policymakers to *"[e]ncourage teacher and leader preparation programs to include the evaluation frameworks in their curricula"* (75 percent or 27 respondents). Similarly, also with 75 percent of participants rating it as a very high priority, the second recommendation (R2) advises state policymakers to *"[i]mprove state budgeting timelines and clarity to allow districts to prevent unnecessary RIFs"* (75 percent or 27 respondents).

- The second highest priority recommendation (R3) advises state policymakers to *"[p]rovide adequate funding to recruit a strong pool of teachers and principals and do not combine it with 'funding games' (whereby resources are moved around but not augmented)"* (65.7 percent or 23 respondents).

- The third highest priority recommendation (R4) advises state policymakers to *"[c]reate stronger alignment between highly qualified teachers (HQT) and endorsements"* (62.9 percent or 22 respondents).

- The fourth highest priority recommendation (R5) advises state policymakers to *"[c]reate a free tool for districts to forecast their funding to allow them to move hiring timelines up to April to allow greater due diligence in hiring and to attract talent before they accept other offers"* (47.2 percent or 17 respondents).

Figure 4. Prioritization of Recommendations on Recruitment and Hiring

R1. Encourage teacher and leader preparation programs to include the evaluation frameworks in their curricula — 2.8% / 22.2% / 75.0%

R2. Improve state budgeting timelines and clarity to allow districts to prevent unnecessary RIFs — 25.0% / 75.0%

R3. Provide adequate funding to recruit a strong pool of teachers and principals and do not combine it with "funding games" (whereby resources are moved around but not augmented) — 11.4% / 22.9% / 65.7%

R4. Create stronger alignment between HQT and endorsements — 11.4% / 25.7% / 62.9%

R5. Create a free tool for districts to forecast their funding to allow them to move hiring timelines up to April to allow greater due diligence in hiring and to attract talent before they accept other offers — 16.7% / 36.1% / 47.2%

R6. Allow "lateral transfer" for RIF'd teachers that have strong evaluation ratings to move to nearby districts (e.g., similar to police department) — 28.6% / 34.3% / 37.1%

R7. Provide districts the opportunity to apply for a waiver from RIFs if their enrollment falls short of good faith forecasts — 14.3% / 51.4% / 34.3%

R8. Create a state work group to strategize on how to improve the pool of candidates entering the profession — 25.0% / 41.7% / 33.3%

R9. Collect data to inform teacher recruitment practices (e.g., how many districts were over/under forecast and would need a waiver, how many RIF'd teachers leave and/or return to the profession — 22.2% / 58.3% / 19.4%

R10. Provide budgeting training to districts to equip them to recruit and hire staff more effectively — 30.6% / 55.6% / 13.9%

■ Not yet ready to address ■ Somewhat ready to address ■ Very ready to address

Next, participants were asked to rate the readiness of the state to address each of the 10 recommendations. Thirty-three survey respondents, or 76.7 percent, answered this question. Figure 5 summarizes the participants' responses. The top three recommendations that policy-to-practice cohort participants said the state was "very ready to address" were the same recommendations participants identified as very high priority, although their rankings are in a slightly different order.

- The recommendation to encourage teacher and leader preparation programs to integrate the evaluation frameworks into their curricula (R1) was one of the highest priority recommendations and the recommendation that participants most thought the state was most ready to address (60.6 percent or 20 respondents).

- The recommendation to improve clarity and timelines for budgeting (R2) was also one of the highest priority recommendations and second recommendation the state was most ready to address (51.5 percent or 17 respondents). The recommendation to improve alignment between HQT and endorsements (R4) was rated as the third highest priority and also received a similarly high endorsement for state readiness to address (51.5 percent or 17 respondents).

- The recommendation to provide adequate funding for recruitment (R3) was rated as the second highest priority and the third recommendation the state was most ready to address (39.4 percent or 13 respondents).
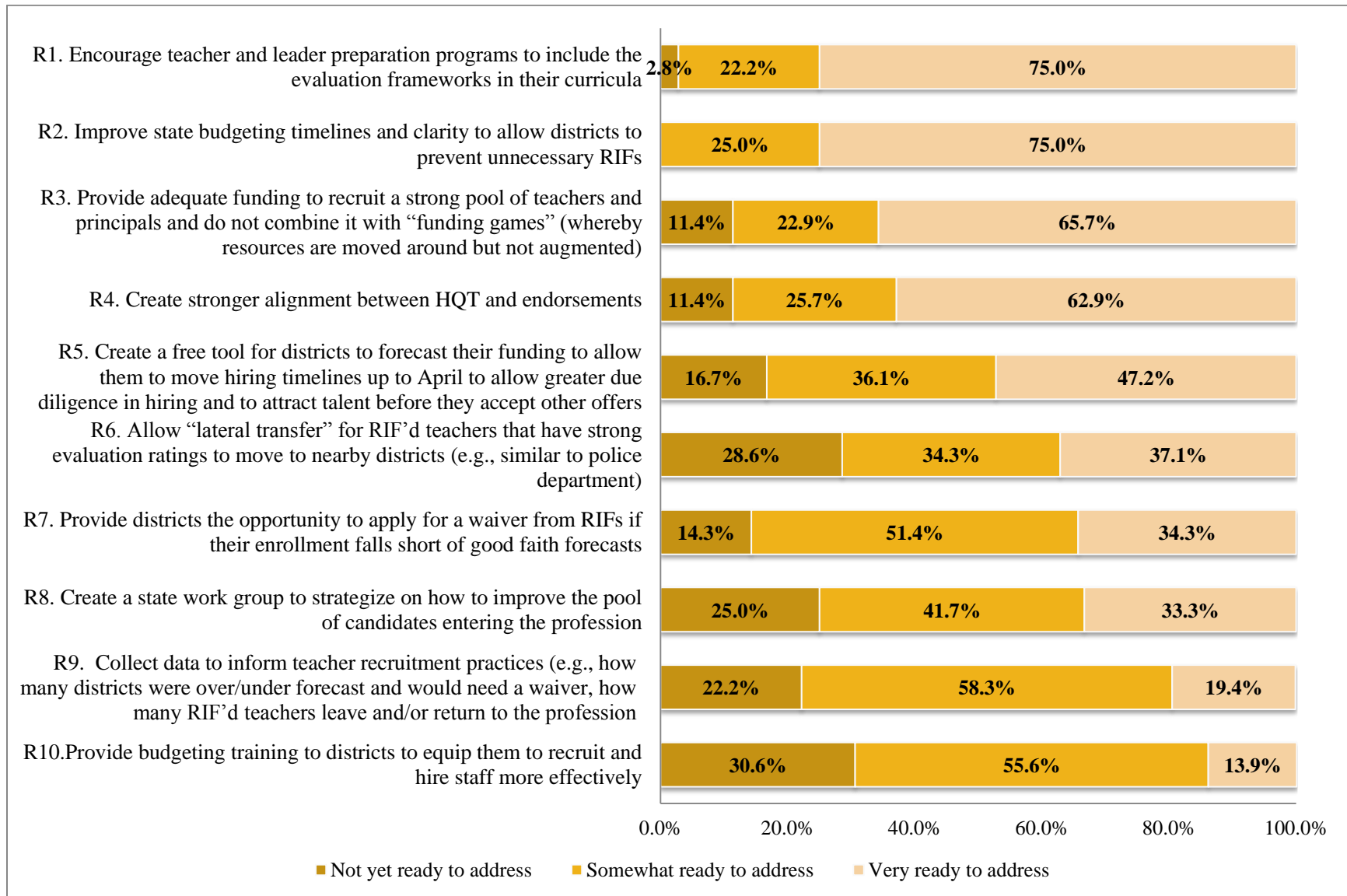
- The recommendation to create a free forecasting tool (R5) was rated as the fourth highest priority recommendation, but only 18.2 percent, or 6 respondents, felt the state was "very ready to address" the recommendation; 51.5 percent, or 17 respondents, felt the state was "somewhat ready to address" the recommendation.

Figure 5. Readiness to Address Recommendations on Recruitment and Hiring

R1. Encourage teacher and leader preparation programs to include the evaluation frameworks in their curricula: 6.1% | 33.3% | 60.6%

R2. Improve state budgeting timelines and clarity to allow districts to prevent unnecessary RIFs: 9.1% | 39.4% | 51.5%

R4. Create stronger alignment between HQT and endorsements: 15.2% | 33.3% | 51.5%

R3. Provide adequate funding to recruit a strong pool of teachers and principals and do not combine it with "funding games" (whereby resources are moved around but not augmented): 18.2% | 42.4% | 39.4%

R8. Create a state work group to strategize on how to improve the pool of candidates entering the profession: 21.2% | 48.5% | 30.3%

R7. Provide districts the opportunity to apply for a waiver from RIFs if their enrollment falls short of good faith forecasts: 30.3% | 39.4% | 30.3%

R6. Allow "lateral transfer" for RIF'd teachers that have strong evaluation ratings to move to nearby districts (e.g., similar to police department): 36.4% | 42.4% | 21.2%

R5. Create a free tool for districts to forecast their funding to allow them to move hiring timelines up to April to allow greater due diligence in hiring and to attract talent before they accept other offers: 30.3% | 51.5% | 18.2%
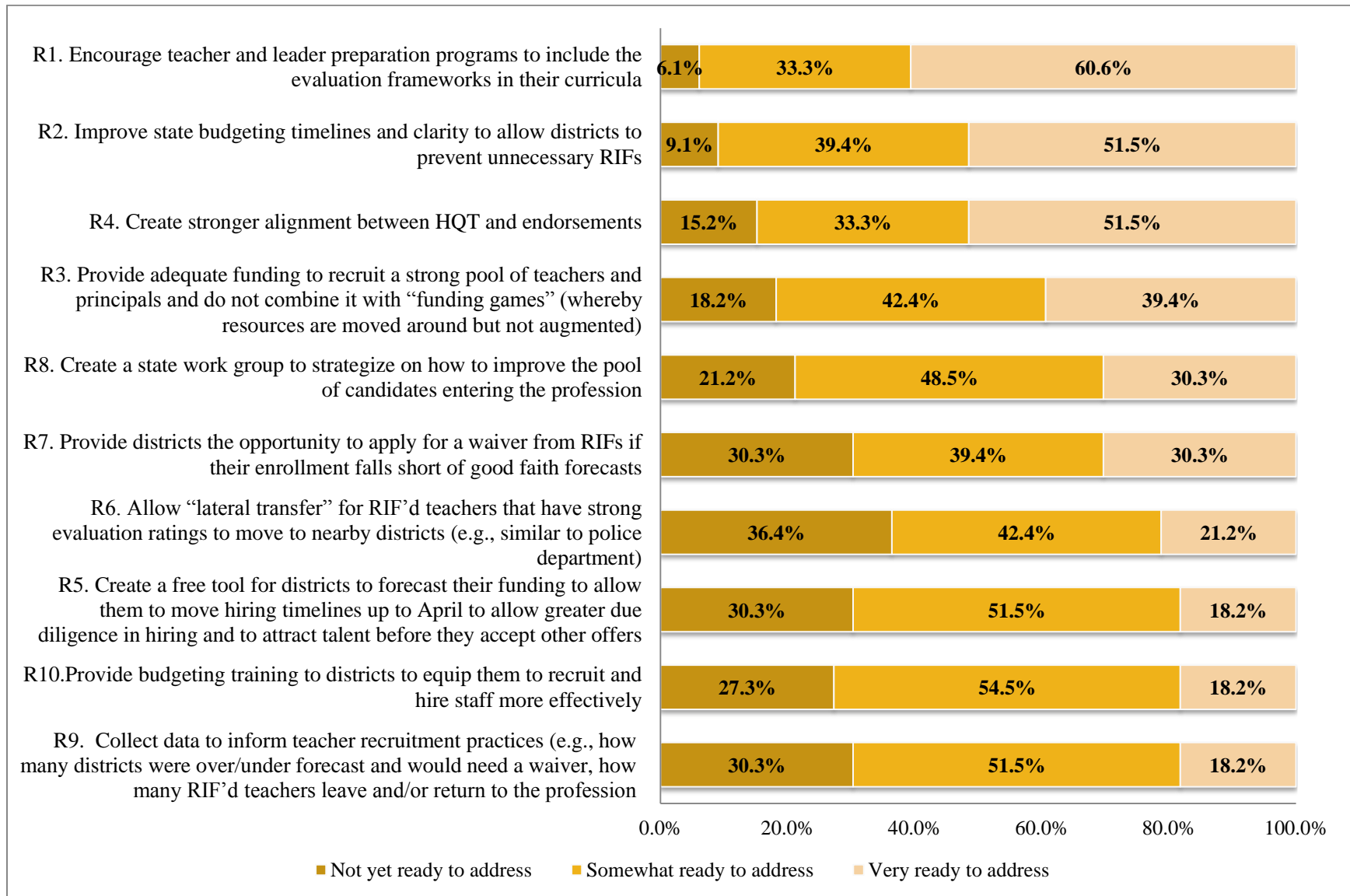
R10. Provide budgeting training to districts to equip them to recruit and hire staff more effectively: 27.3% | 54.5% | 18.2%

R9. Collect data to inform teacher recruitment practices (e.g., how many districts were over/under forecast and would need a waiver, how many RIF'd teachers leave and/or return to the profession: 30.3% | 51.5% | 18.2%

Legend: ■ Not yet ready to address   ■ Somewhat ready to address   ■ Very ready to address
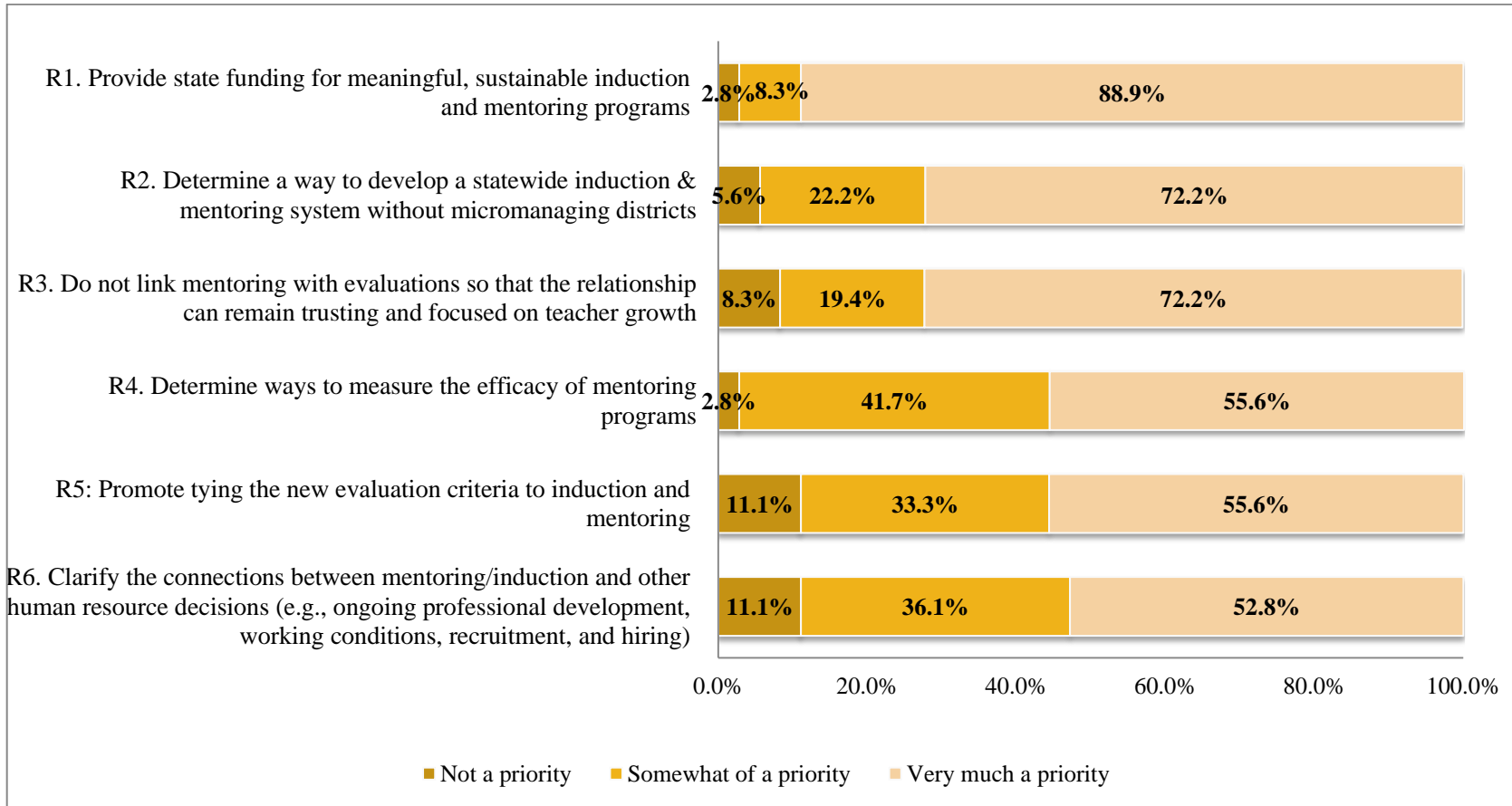
Recommendations for Mentoring and Induction. The policy-to-practice cohorts developed six recommendations on this topic area. Respondents were asked to rate the priority of each recommendation. Thirty-six survey respondents, or 83.7 percent, answered this question. Figure 6 presents each recommendation in order of priority from highest to lowest. The top four recommendations that policy-to-practice cohort participants identified as "very much a priority" address funding and statewide support, as well as the question of whether or not to link mentoring and induction with evaluation results.

- The highest priority recommendation (R1) advises state policymakers to *"[p]rovide state funding for meaningful, sustainable induction and mentoring programs"* (88.9 percent or 32 respondents).

- The second highest priority recommendation (R2) advises state policymakers to *"[d]etermine a way to develop a statewide induction & mentoring system without micromanaging districts"* (72.2 percent or 26 respondents). Similarly, with 72.2 percent of respondents rating it as a very high priority, the third-ranked recommendation (R3) advises state policymakers to *"[n]ot link mentoring with evaluations so that the relationship can remain trusting and focused on teacher growth"* (26 respondents).

- In direct contrast, the fourth highest priority recommendation (R4) advises state policymakers to *"[d]etermine ways to measure the efficacy of mentoring programs"* (55.6 percent or 20 respondents). Similarly, with 55.6 percent of respondents rating it as a very high priority, the fifth-ranked recommendation (R5) advises state policymakers to *"[p]romote tying the new evaluation criteria to induction and mentoring"*[5] (20 respondents).

---

[5] At first glance, recommendations 3 and 5 may seem at odds; however, the participant discussions on these recommendations suggest that recommendation 3 is aimed more squarely at ensuring mentors are not also serving as evaluators whereas recommendation 5 is focused on ensuring the new evaluation *criteria* are included as part of induction and mentoring programs.

Figure 6. Prioritization of Recommendations on Mentoring and Induction

R1. Provide state funding for meaningful, sustainable induction and mentoring programs — Not a priority 2.8%, Somewhat of a priority 8.3%, Very much a priority 88.9%

R2. Determine a way to develop a statewide induction & mentoring system without micromanaging districts — Not a priority 5.6%, Somewhat of a priority 22.2%, Very much a priority 72.2%

R3. Do not link mentoring with evaluations so that the relationship can remain trusting and focused on teacher growth — Not a priority 8.3%, Somewhat of a priority 19.4%, Very much a priority 72.2%

R4. Determine ways to measure the efficacy of mentoring programs — Not a priority 2.8%, Somewhat of a priority 41.7%, Very much a priority 55.6%

R5: Promote tying the new evaluation criteria to induction and mentoring — Not a priority 11.1%, Somewhat of a priority 33.3%, Very much a priority 55.6%

R6. Clarify the connections between mentoring/induction and other human resource decisions (e.g., ongoing professional development, working conditions, recruitment, and hiring) — Not a priority 11.1%, Somewhat of a priority 36.1%, Very much a priority 52.8%

■ Not a priority   ■ Somewhat of a priority   ■ Very much a priority

Next, participants were asked to rate the readiness of the state to address each of the 10 recommendations. Thirty-four survey respondents, or 79 percent, answered this question. Figure 7 summarizes the participants' responses. The top four recommendations that policy-to-practice cohort participants said the state was "very ready to address" were the same recommendations participants identified as very high priority, although their rankings are in a slightly different order.

- The recommendation to provide state funding for mentoring and induction (R1) was the highest priority recommendation and the recommendation that participants most thought the state was most ready to address (70.6 percent or 24 respondents).

- The recommendation to not link evaluations with mentoring and induction programs (R3) was one of the second highest priority recommendations and the second recommendation the state was most ready to address (58.8 percent or 20 respondents).

- The recommendation to develop a statewide system for mentoring and induction that does not micromanage districts (R2) was one of the second highest priority recommendations and the third recommendation the state was most ready to address (52.9 percent or 18 respondents).

- The recommendation to promote tying evaluation criteria with mentoring and induction (R5) was rated as the third highest priority recommendation and the fourth recommendation the state was most ready to address (44.1 percent or 15 respondents).

Figure 7. Readiness to Address Recommendations on Mentoring and Induction

Recommendation	Not yet ready to address	Somewhat ready to address	Very ready to address
R1. Provide state funding for meaningful, sustainable induction and mentoring programs	11.8%	17.6%	70.6%
R3. Do not link mentoring with evaluations so that the relationship can remain trusting and focused on teacher growth	26.5%	14.7%	58.8%
R2. Determine a way to develop a statewide induction & mentoring system without micromanaging districts	14.7%	32.4%	52.9%
R4. Promote tying the new evaluation criteria to induction and mentoring	26.5%	29.4%	44.1%
R4. Determine ways to measure the efficacy of mentoring programs	11.8%	50.0%	38.2%
R6. Clarify the connections between mentoring/induction and other human resource decisions (e.g., ongoing professional development,…	20.6%	41.2%	38.2%

■ Not yet ready to address   ■ Somewhat ready to address   ■ Very ready to address

Recommendations for Principal Authority over Teacher Placement. The policy-to-practice cohorts developed five recommendations on this topic area. Respondents were asked to rate the priority of each recommendation. Thirty-five survey respondents, or 81.4 percent, answered this question. Figure 8 presents each recommendation in order of priority from highest to lowest. The top three recommendations that policy-to-practice cohort participants identified as "very much a priority" focus on both defining principal authority and retaining local control and bargaining over teacher placement.

- The highest priority recommendation (R1) advises state policymakers to *"[c]reate a legislative requirement controlling/restricting the latitude of principal authority; this should remain locally bargained"* (70.6 percent or 24 respondents).

- The second highest priority recommendation (R2) advises state policymakers that *"[a]s long as the details can be locally determined, provide principals a voice in placement of teachers into their building"* (68.6 percent or 24 respondents).

- The third highest priority recommendation (R3) advises state policymakers to *"...not make placement decisions based on evaluations yet; wait until TPEP has been used for some time consistently, reliably and effectively* (60 percent or 21 respondents).
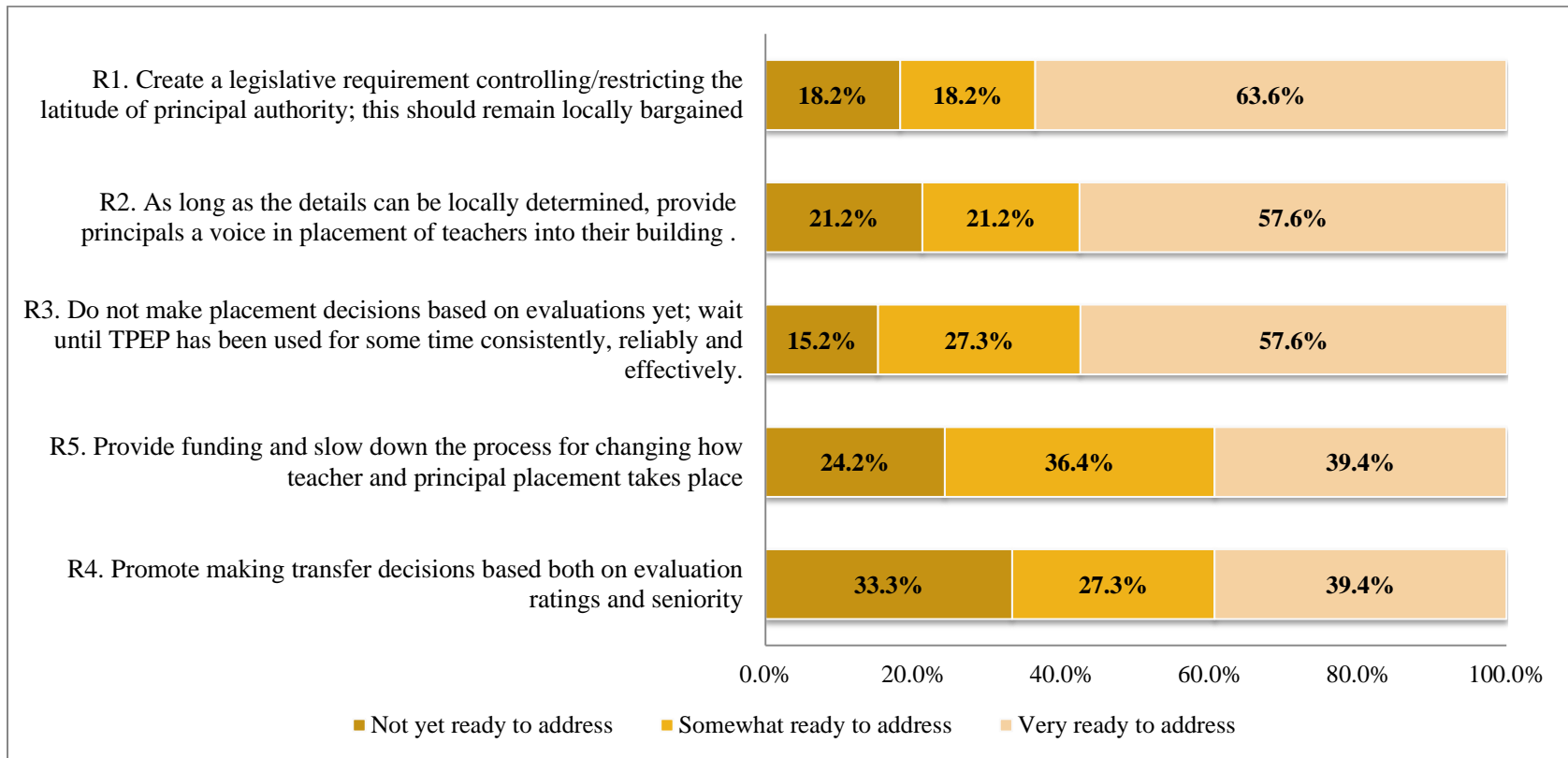
Figure 8. Prioritization of Recommendations on Placement (Principal Authority)

Recommendation	Not a priority	Somewhat of a priority	Very much a priority
R1. Create a legislative requirement controlling/restricting the latitude of principal authority; this should remain locally bargained	8.8%	20.6%	70.6%
R2. As long as the details can be locally determined, provide principals a voice in placement of teachers into their building	17.1%	14.3%	68.6%
R3. Do not make placement decisions based on evaluations yet; wait until TPEP has been used for some time consistently, reliably and effectively.	11.4%	28.6%	60.0%
R4. Promote making transfer decisions based both on evaluation ratings and seniority	26.5%	29.4%	44.1%
R5. Provide funding and slow down the process for changing how teacher and principal placement takes place	17.1%	57.1%	25.7%

■ Not a priority ■ Somewhat of a priority ■ Very much a priority

Next, participants were asked to rate the readiness of the state to address each of the five recommendations. Thirty-three survey respondents, or 76.7 percent, answered this question. Figure 9 summarizes the participants' responses. The top three recommendations that policy-to-practice cohort participants said the state was "very ready to address" were the same recommendations participants identified as very high priority.

- The recommendation to create restrictions on principal authority over teacher placement (R1) was the highest priority recommendation and the recommendation that participants most thought the state was most ready to address (63.6 percent or 21 respondents).

- The recommendation to provide principals a voice in teacher placement within locally bargained agreements (R2) was the second highest priority recommendation and second recommendation that participants felt the state was most ready to address (57.6 percent or 19 respondents). The third highest priority recommendation to delay linking placement and evaluations (R3) was also rated by 57.6 percent of respondents as the recommendation the state was most ready to address.
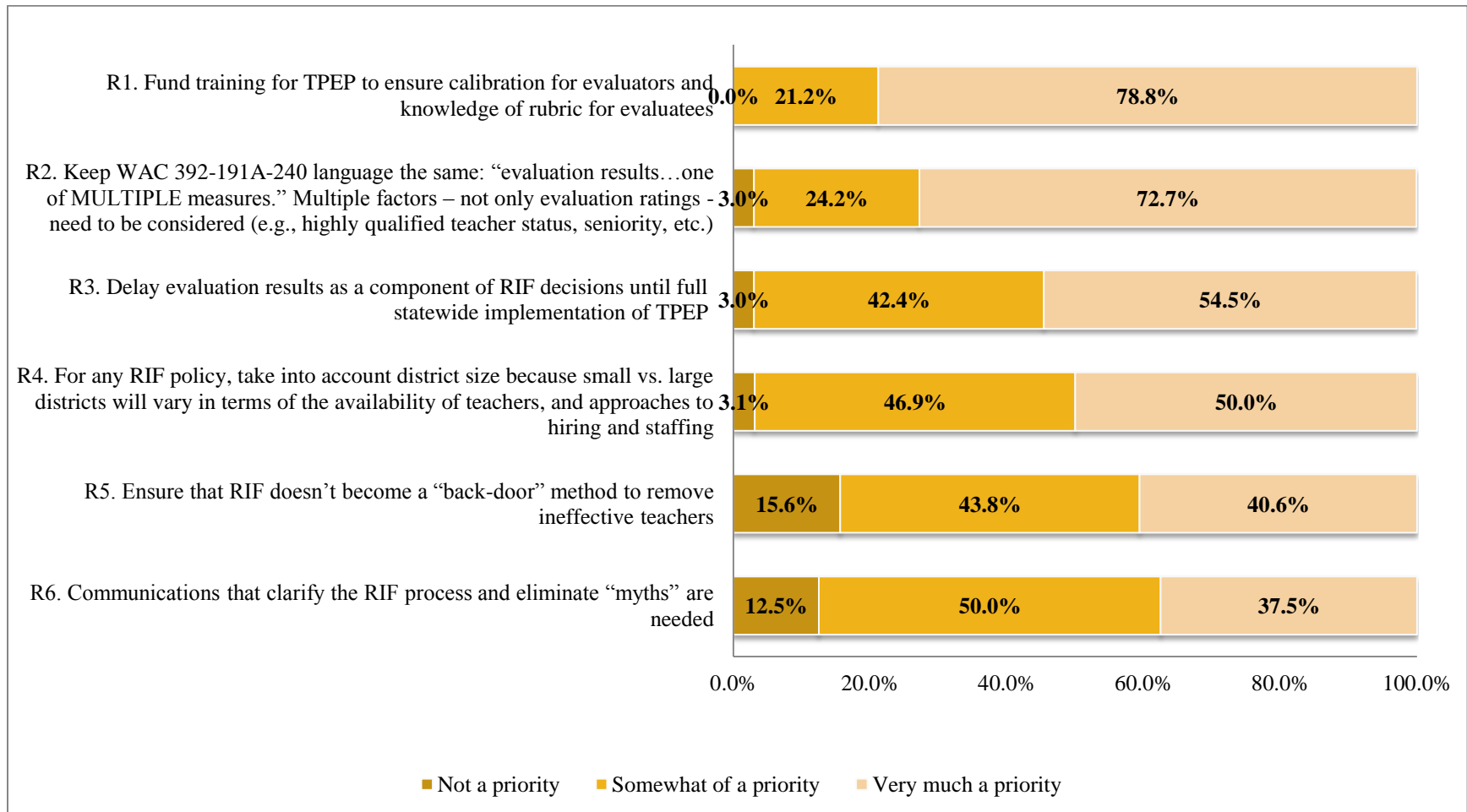
Figure 9. Readiness to Address Recommendations on Placement (Principal Authority)

R1. Create a legislative requirement controlling/restricting the latitude of principal authority; this should remain locally bargained: 18.2% | 18.2% | 63.6%

R2. As long as the details can be locally determined, provide principals a voice in placement of teachers into their building .: 21.2% | 21.2% | 57.6%

R3. Do not make placement decisions based on evaluations yet; wait until TPEP has been used for some time consistently, reliably and effectively.: 15.2% | 27.3% | 57.6%

R5. Provide funding and slow down the process for changing how teacher and principal placement takes place: 24.2% | 36.4% | 39.4%

R4. Promote making transfer decisions based both on evaluation ratings and seniority: 33.3% | 27.3% | 39.4%

■ Not yet ready to address ■ Somewhat ready to address ■ Very ready to address

Recommendations for Reduction in Force (RIF). The policy-to-practice cohorts developed six recommendations on this topic area. Respondents were asked to rate the priority of each recommendation. Thirty-three survey respondents, or 76.7 percent, answered this question. Figure 10 presents each recommendation in order of priority from highest to lowest. The top four recommendations that policy-to-practice cohort participants identified as "very much a priority" focus on funding, the importance of multiple measures, and appropriate timing and considerations for implementation.

- The highest priority recommendation (R1) advises state policymakers to *"[f]und training for TPEP to ensure calibration for evaluators and knowledge of rubric for evaluatees"* (78.8 percent or 26 respondents).

- The second highest priority recommendation (R2) advises state policymakers to *"[k]eep WAC 392-191A-240 language the same: "evaluation results…one of MULTIPLE measures." Multiple factors – not only teacher or principal evaluation ratings – need to be considered (e.g., highly qualified teacher status, seniority, etc."* (72.7 percent or 24 respondents).

- The third highest priority recommendation (R3) advises state policymakers to *"[d]elay evaluation results as a component of RIF decisions until full statewide implementation of TPEP"* (54.5 percent or 18 respondents).

- The fourth highest priority recommendation (R4) advises state policymakers *"[f]or any RIF policy, take into account district size because small vs. large districts will vary in terms of the availability of teachers, and approaches to hiring and staffing"* (50 percent or 16 respondents).
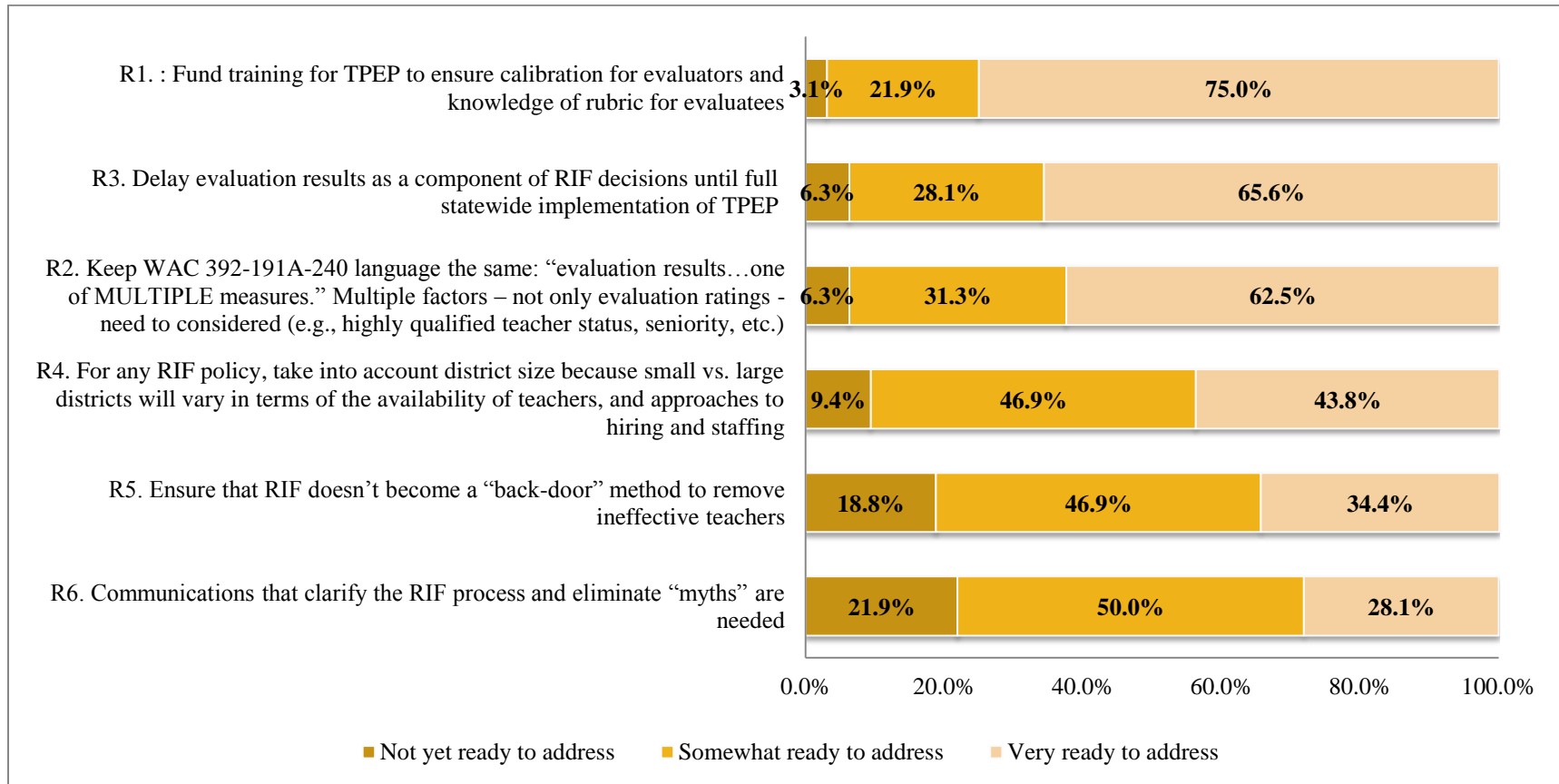
Figure 10. Prioritization of Recommendations on Reduction in Force

R1. Fund training for TPEP to ensure calibration for evaluators and knowledge of rubric for evaluatees: **0.0%** | **21.2%** | **78.8%**

R2. Keep WAC 392-191A-240 language the same: "evaluation results…one of MULTIPLE measures." Multiple factors – not only evaluation ratings - need to be considered (e.g., highly qualified teacher status, seniority, etc.): **3.0%** | **24.2%** | **72.7%**

R3. Delay evaluation results as a component of RIF decisions until full statewide implementation of TPEP: **3.0%** | **42.4%** | **54.5%**

R4. For any RIF policy, take into account district size because small vs. large districts will vary in terms of the availability of teachers, and approaches to hiring and staffing: **3.1%** | **46.9%** | **50.0%**

R5. Ensure that RIF doesn't become a "back-door" method to remove ineffective teachers: **15.6%** | **43.8%** | **40.6%**

R6. Communications that clarify the RIF process and eliminate "myths" are needed: **12.5%** | **50.0%** | **37.5%**

■ Not a priority ■ Somewhat of a priority ■ Very much a priority

Next, participants were asked to rate the readiness of the state to address each of the six recommendations. Thirty-two survey respondents, or 74.4 percent, answered this question. Figure 11 summarizes the participants' responses. The top three recommendations that policy-to-practice cohort participants said the state was "very ready to address" were the same recommendations participants identified as very high priority, although in a slightly different order.

- The recommendation to fund evaluator calibration training and evaluatee training (R1) was the highest priority recommendation and the recommendation that participants most thought the state was most ready to address (75 percent or 24 respondents).

- The recommendation to delay the use of evaluation results in RIF decisions until full statewide implementation (R3) was the third highest priority recommendation and second recommendation that participants felt the state was most ready to address (65.6 percent or 21 respondents).

- The second highest priority recommendation to keep WAC language referring to the use of multiple measures in RIF decisions (R2) was rated by 62.5 percent, or 20 respondents, as the recommendation the state was most ready to address.

Figure 11. Readiness to Address Recommendations on Reduction in Force

R1. : Fund training for TPEP to ensure calibration for evaluators and knowledge of rubric for evaluatees: 3.1% | 21.9% | 75.0%

R3. Delay evaluation results as a component of RIF decisions until full statewide implementation of TPEP: 6.3% | 28.1% | 65.6%

R2. Keep WAC 392-191A-240 language the same: "evaluation results…one of MULTIPLE measures." Multiple factors – not only evaluation ratings - need to considered (e.g., highly qualified teacher status, seniority, etc.): 6.3% | 31.3% | 62.5%

R4. For any RIF policy, take into account district size because small vs. large districts will vary in terms of the availability of teachers, and approaches to hiring and staffing: 9.4% | 46.9% | 43.8%

R5. Ensure that RIF doesn't become a "back-door" method to remove ineffective teachers: 18.8% | 46.9% | 34.4%

R6. Communications that clarify the RIF process and eliminate "myths" are needed: 21.9% | 50.0% | 28.1%

■ Not yet ready to address ■ Somewhat ready to address ■ Very ready to address

Policy-to-Practice Cohort Summary

The policy-to-practice cohorts represent a unique, ongoing opportunity to listen to the expertise, reflections, and recommendations of educators working "in the trenches" implementing evaluation reform. Although the advice and recommendations generated in each cohort is fairly specific to each topic area, across all three policy-to-practice cohorts several common themes emerged including the following:

- **Funding:** in every topic area, cohort participants repeatedly voiced their concerns about the practical realities of implementing and sustaining new evaluation policies and systems without access to sufficient funds. As pilot districts, a number of participants remarked that it was difficult to imagine how a district without access to the funding and support TPEP sites received could successfully implement a new system. Moreover, participants consistently pointed out the importance of thinking carefully about funding and support for new requirements around linking evaluation results to human capital management decisions. Participants consistently expressed great hopes in the potential and possibilities that the new evaluation requirements pose, but reminded policymakers that sufficient funding and support are prerequisites to realizing those potentialities.

- **Flexibility:** districts and schools in Washington vary widely in type, context, capacity, and needs. Participants lauded the current statewide approach that aims to give districts flexibility in selecting a model and design that best fits their unique needs. Across all cohort topics, participants encouraged policymakers to continue to focus on prioritizing flexibility for districts as they move forward with formulating regulations and guidance for professional development and human capital management decisions.

- **Alignment and coherence:** with the incredible number of initiatives, programs, and requirements that schools and districts are tasked with implementing, the potential for evaluation system reform to become just another layer of "compliance activities" is a very real concern for participants. Specifically, participants encouraged districts to thinking about how curriculum, instruction, goal-setting, and assessment can be aligned as part of the evaluation system reform, rather than merely adding it on as another loosely connected system. Similarly, in discussing human capital management, participants were very vocal about the need for state-level policymakers to create alignment and coherence across preparation, certification, licensure, and evaluation. Currently, the burden on new teachers for demonstrating their proficiency is heavy and it's critical that the new evaluation systems serve to reinforce and work in conjunction with other requirements.

- **Communication:** clarity in messaging from the legislature down to the school-level are crucial for building educator trust and ensuring the culture shifts envisioned by the evaluation system reform are actually possible. Participants strongly recommended further support from the state in crafting consistent messages and utilizing effective communication methods to ensure the widest possible understanding and engagement from the field at large.

Communications Toolkit

Across each of the topic areas (student growth, professional learning, and human resource decisions) discussed above, pilot districts regularly mentioned the importance of clear and consistent information and messaging at all levels. Pilot districts reported that the unexpected shift to the policy in ESSB 5895 created extensive anxiety, frustration, and concern among educators in their communities. Districts also emphasized the importance of a single coherent message from both OSPI and the Washington state legislature about the purposes of the evaluation reforms, the new requirements, and about how evaluation results will be used by both the state and districts. After holding constituent discussions with educators in their districts, several pilot districts discovered that many teachers remained uninformed on key evaluation questions, even after focused communication efforts as part of the pilot.

Districts have a critical role to play in this process as an important link between state-level communication and individual educators. As indicated in the statewide survey results, one third of teachers are not well informed about state and district evaluation reforms. Moreover, one key area that educators identified as an area of need was access to information and lessons-learned from the TPEP pilot districts.

Resources developed through the Year 3 scope of work are designed to assist Washington school districts with strategies to better communicate about both statewide policy changes and new district-designed evaluation systems. With feedback and support from OSPI, TPEP pilot districts, and TPEP Steering Committee members, AIR created a communications toolkit that a range of Washington stakeholders can use as part of their communications efforts. Each piece of the toolkit is described below.

Capturing the Learning Points: How Teacher/Principal Evaluation Pilots Reformed Educator Evaluation

Capturing the Learning Points: How Teacher/Principal Evaluation Pilots Reformed Educator Evaluation includes a series of five short, user-friendly publications intended to support school districts statewide by providing details on both the practical aspects of adopting a new evaluation system and pilot districts' ideas and insights on critical topic areas. The five publications include:

- Early Decision-making and Engaging Stakeholders
- Developing Observations that Result in Professional Conversations
- Measuring Student Growth
- Time-saving Strategies for Principals
- Implementing New Evaluations: A Teacher's Perspective

Capturing the Learning Points publications help Washington districts to learn from pilot districts' experiences, ease anxiety around evaluation reform, and make the most of the reforms in achieving their district's mission. Each publication highlights the following elements:

- The steps taken at each stage of the reform process, including the individuals involved and other logistics

- The issues and decision points that arose in the process of teacher and/or principal evaluation redesign

- Tools or resources that were accessed or developed to aid the transition to a new evaluation system

- Lessons learned, unanticipated implementation issues that arose, and other advice for districts at an earlier stage of reform

District Communication Guide

To assist districts more fully in communicating with stakeholders about the new system and how it can further the district's mission, AIR created a district communication guide. The guide is designed to help Washington school districts determine the materials that should be posted on their websites and to identify the communication strategies and tactics that are likely to aid smooth evaluation reform implementation in their context. The guide is intended to help districts communicate consistent and appropriate messages about these reforms to each of their identified audiences.

Educator Evaluation Infographic

To assist educators in visualizing the "big picture" of educator evaluation reforms in Washington, AIR and OSPI created a one-page, user-friendly "info-graphic." An info-graphic is a fun and interesting way of visually presenting a set of data or information. The "Teacher and Principal Evaluation" info-graphic summarizes the core principles, requirements, and timelines for teacher and principal evaluation reforms in an engaging manner that can quickly and easily be examined by a range of stakeholders.

Online Professional Learning Modules

To support the requirements of ESB 5895, Section 5, AIR developed seven online modules to orient evaluators and educators to the new system. The modules are intended for use by school leadership teams to facilitate professional development sessions for principals, administrators, and teachers. Each module is designed for presentation in two- to three- hour training sessions, but the materials are flexible enough to be combined or broken apart as best meets district needs. The materials can be delivered in person or through recorded webinars and are designed for use in a self-directed manner. The purpose of the modules is to support school leadership teams in developing a common understanding of the educator evaluation legislation and encourage shared language about educator evaluation across the state of Washington. The seven modules cover the following topics:

- **Introduction to Educator Evaluation in Washington.** The basics of educator evaluation reforms, the evaluation criteria for teachers and principals, the four-level rating system, the state and local decision matrix, and a preview of the remaining modules.

- **Using Instructional and Leadership Frameworks in Educator Evaluation.** An orientation to the components of instructional and leadership frameworks, how they are different from previous evaluation tools, and how they support identification of practice across a continuum. This module will provide a "jumpstart" into the three instructional and two leadership frameworks.

- **Preparing and Applying Formative Multiple Measures of Performance: An Introduction to Self-Assessment, Goal Setting, and Criterion Scoring.** An overview of the types of measures required and supported by RCW 28A.405.100 and WAC 392-191A, the differences between measures and evidence, how to move beyond an observation-only evaluation system, and the benefits to the types of measures used in educator evaluation. This module includes criterion scoring guidance that has been informed by the instructional and leadership framework authors.

- **Including Student Growth in Educator Evaluation.** An overview of goal setting for student growth, selecting classroom-based, school-based, district-based, and state-based tools, and using student learning data in educator evaluation. This module will offer a process for establishing student growth goals, examples of student growth goals, and a process for determining the change in student achievement between two points in time.

- **Conducting High-Quality Observations and Maximizing Rater Agreement.** An overview of high-quality observation practices with special emphasis on collecting evidence, strategies for maximizing rater agreement, and strategies for districts and school administrators to learn about and plan for maximizing rater agreement.

- **Providing High-Quality Feedback for Continuous Professional Growth and Development.** An overview of examples of, and protocols for, how to provide feedback to teachers and leaders so that they continue to grow and improve in their practice; how to engage faculty in these conversations; and strategies for connecting professional development planning with evaluation outcomes.

- **Combining Multiple Measures into a Summative Rating.** Although an evaluation rating is often viewed as a measure of a single point in time, it is actually the culmination of a comprehensive process of self-assessment, goal-setting, plan implementation, dialogue, and reflection that unfolds during several months. This module provides an understanding of how to assess practice using multiple types of evidence and performance rubrics in a thoughtful, comprehensive, reliable manner and follow the process from OSPI to create an overall performance rating.

Each module consists of a downloadable package of materials meant to span a three- to four-hour session plus additional application activities. Each package includes the following:

- A PowerPoint presentation to guide and pace participants' learning

- An annotated facilitator's guide to support facilitators in presenting and communicating the material

- A participant handouts packet that contains all materials needed to support the engagement activities in the module

The modules are organized into a four-part structure to help facilitators and participants pace the content appropriately. The four segments of each module are as follows:

- **Connecting**—Builds community, prepares the team for learning, and links to prior knowledge, other modules, and current work

- **Learning**—Describes key concepts and highlights various implementation scenarios; supports teams in applying knowledge and sharing ideas

- **Implementing**—Supports teams in problem solving and planning next steps for schools and districts

- **Reflecting**—Engages participants in providing feedback, reflecting on learning, and closing the session

OSPI has placed each module on the TPEP website and on OSPI's iTunes U account to facilitate easy access and use by educators across the state.

Educator Evaluation Annotated Bibliographies

To support districts, particularly district design committees, as they move through the planning and implementation process, AIR and OSPI created a series of annotated bibliographies on six key educator evaluation topics. Each bibliography contains four to five references, and short summaries of each reference. The topics include the following:

- Rater agreement
- Goal setting linked to student outcomes and test scores
- Using frameworks, rubrics, and common language to improve student learning
- Student growth measures
- Evaluation system implementation
- Human resource connections

References

Alexander, K. L, Entwisle, D. R., & Olson, L. S., (2007). Lasting consequences of the summer learning gap. *American Sociological Review*. *72* (April). Retrieved from http://www.nayre.org/Summer%20Learning%20Gap.pdf

Brown-Sims,M., Clayton, T., Chen, J., and Brandt, C. (2013). *Washington's 2012 state of the state educator survey report: Summary of key findings.* Washington, D.C.: American Institutes for Research. Retrieved from http://tpep-wa.org/wp-content/uploads/1408_WA%20TPEP_Yr3_StatewideSurveyReport_ed_AllEducators_FINAL_03%2022%2013.pdf

Center on Great Teachers and Leaders. (2013). *Databases on state teacher and principal evaluation policies.* Washington, D.C.: American Institutes for Research. Retrieved from http://resource.tqsource.org/stateevaldb/

Charlotte-Mecklenburg Schools. (2009). *Strategic staffing: 'The moral thing to do.'* Charlotte, NC: Author. Retrieved from http://www.cms.k12.nc.us/superintendent/White%20Papers/Strategic%20Staffing.pdf

Community Training and Assistance Center (CTAC). (2004). *Catalyst for change: Pay for performance in Denver.* Final report. Retrieved from http://www.ctacusa.com/PDFs/Rpt-CatalystChangeFull-2004.pdf

Community Training and Assistance Center (CTAC). (2013). *It's more than money: Teacher incentive fund—leadership for educators' advanced performance*. Charlotte-Mecklenburg Schools. Retrieved from http://www.ctacusa.com/PDFs/MoreThanMoney-report.pdf

Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology, 80*(5), 565-579.

Curtis, R. & Weiner, R. (2012). *Means to an end: Guide to developing teacher evaluation systems that support growth and development*. Washington D.C.: Aspen Institute. Retrieved from http://www.aspendrl.org/portal/browse/DocumentDetail?documentId=1008&download

Daly, T. (2011, October). *Transforming teacher evaluations* [Webinar]. Presentation of The New Teacher Project at the National Center for Teacher Effectiveness: Supporting Rater Accuracy and Consistency in Classroom Observation. Retrieved from http://www.gse.harvard.edu/ncte/news/NCTE_Webinar_TNTP_Powerpoint.pdf

Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching.* Stanford, CA: Stanford Center for Opportunity Policy in Education. Retrieved from http://edpolicy.stanford.edu/sites/default/files/publications/creating-comprehensive-system-evaluating-and-supporting-effective-teaching.pdf

Engrossed Second Substitute Senate Bill 6696, 61st Legislature, Regular Session. (2010). Retrieved from http://apps.leg.wa.gov/documents/billdocs/2009-10/Pdf/Bills/Senate Passed Legislature/6696-S2.PL.pdf

Engrossed Substitute Senate Bill 5895, 62nd Legislature, Regular Session. (2012). Retrieved from http://apps.leg.wa.gov/documents/billdocs/2011-12/Pdf/Bills/Senate%20Bills/5895-S.E.pdf

Franck, V., Kelliher, K., & Varghese, P. (2011). *Teacher quality roadmap: Improving policies and practices in Springfield, Massachusetts*. Washington, DC: National Council on Teacher Quality. Retrieved from http://mbae.org/uploads/NCTQ_Springfield.pdf

Fry, R., & Ramsdell, R. (2011, October). *Supporting rater accuracy and consistency in classroom observation* [Webinar]. Presentation of Cambridge Education at the National Center for Teacher Effectiveness: Supporting Rater Accuracy and Consistency in Classroom Observation. Retrieved from http://www.gse.harvard.edu/ncte/news/NCTE_Webinar_Cambridge_Powerpoint.pdf

Glazerman, S., Protik, A., Druch, J., & Saftor, N. (2012). *Moving high-performing teachers: Implementation of transfer incentives in seven districts*. National Center for Education Evaluation and Regional Assistance. Retrieved from http://www.mathematica-mpr.com/publications/pdfs/education/TTI_fnlrpt.pdf

Goe, L. (2008). *Using value-added models to identify and support highly effective teachers* (Key Issue). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://www2.tqsource.org/strategies/het/UsingValueAddedModels.pdf

Goe, L., Bell, C., & Little, O. (2008). *Value-added models and other approaches to evaluating teacher effectiveness: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality.Goldhaber, D. (2011). *A worm in the apple? The implications of seniority-based teacher layoffs.* Washington, DC: American Enterprise Institute for Public Policy Research. Retrieved from http://www.aei.org/files/2011/01/13/WormintheApple.pdf

Goldhaber, D. and Theobald, R. (2012). *Do different value-added models tell us the same things?* Carnegie Knowledge Network. Retrieved from: http://www.carnegieknowledgenetwork.org/briefs/value-added/different-growth-models/

Graham, M., Milanowski, A., & Miller, J. (2012). *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*. Washington, D.C., Center for Educator Compensation Reform. Retrieved from http://cecr.ed.gov/pdfs/Inter_Rater.pdf

Hamman, W. R., Beaubien, J. M., & Holt, R. W. (1999). *Evaluating instructor/evaluator inter-rater reliability from performance database information.* Proceedings of the Tenth International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University Press.

Harris, D. & Anderson, A. (2012). *Bias of public sector worker performance monitoring: Theory and empirical evidence from middle school teachers.* Association for Public Policy Analysis & Management Annual Conference. Retrieved from http://appam.confex.com/appam/2012/webprogram/Paper2262.html

Hassel, E. A., Kowal, J., Ableidinger, J., & Hassel, B. (2011). *Teacher tenure reform: Applying lessons from the civil service and higher education.* Chapel Hill, NC: Public Impact.

Jackson, C. K. (2012), *Teacher quality at the high school level: The importance of accounting for tracks.* NBER Working Paper No. 17722. Retrieved from http://works.bepress.com/c_kirabo_jackson/22/

Kane, T., and Darling-Hammond, L. (2012). Should student test scores be used to evaluate teachers? *Wall Street Journal*. June 24, 2012. Retrieved from http://online.wsj.com/article/SB10001424052702304723304577366023832205042.html

Lachlan-Haché, L., Cushing, E., and Bivona, L. (2012). *Student learning objectives as measures of educator effectiveness: The basics*. Washington, DC: American Institutes for Research. Retrieved from: http://educatortalent.airprojects.org/inc/docs/SLOs_Measures_of_Educator_Effectiveness.pdf

Lemke, M., Thomsen, K., Wayne, A., and Birman, B. (2012). *Providing effective teachers for all students: Examples from five districts.* U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Policy and Program Studies Service. Washington, D.C. Retrieved from http://www2.ed.gov/rschstat/eval/teaching/providing-effective-teachers/report.pdf

Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to the raters? *Language Testing, 19*, 246–277.

McCaffrey, D. F., Koretz, D., Lockwood, J. R., & Hamilton, L. S. (2005). *Evaluating value-added models for teacher accountability.* Santa Monica, CA: RAND Corporation.

McClellan, C., Atkinson, M., & Danielson, C. (2012). *Teacher evaluator training and certification: Lessons learned from the Measures of Effective Teaching project* (paper in the Practitioner Series for Teacher Evaluation). San Francisco: Teachscape. Retrieved from http://www.teachscape.com/binaries/content/assets/teachscape-marketing-website/resources/march_13whitepaperteacherevaluatortraining.pdf

McGuinn, P. (2010). *Ringing the bell for K-12 teacher tenure reform.* Washington, DC: Center for American Progress. Retrieved from http://www.americanprogress.org/wp-content/uploads/issues/2010/02/pdf/teacher_tenure.pdf

Melchers, K. G., Lienhardt, N., Von Aarburg, M., & Kleinmann, M. (2011). Is more structure really better? A comparison of frame-of-reference training and descriptively anchored rating scales to improve interviewers' rating quality. *Personnel Psychology, 64*, 53–87.

MET Project. (2013). *Ensuring fare and reliable measures of effective teaching: Culminating findings from the MET project's three-year study.* Bill and Melinda Gates Foundation. Seattle, WA. Retrieved from http://metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf

National Council on Teacher Quality (2010a). *2010 state teacher policy yearbook: Blueprint for change: National summary.* Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/stpy09/updates/docs/stpy_national.pdf

National Council on Teacher Quality (2010b). *Teacher layoffs: Rethinking "last hired – first fired" policies.* Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/p/docs/nctq_dc_layoffs.pdf

National Council on Teacher Quality (2012). *2012 state teacher policy yearbook: Brief Area 3: Identifying effective teachers.* Washington, DC: National Council on Teacher Quality. Retrieved from http://www.nctq.org/p/publications/docs/Updated_NCTQ_State of the States 2012_Teacher Effectiveness Policies.pdf

Newton, X., Darling-Hammond, L., Haertel, E., & Thomas, E. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Educational Policy Analysis Archives, 18*(23). Retrieved from http://epaa.asu.edu/ojs/article/view/810

Prince, C. D., Schuermann, P. J., Guthrie, J. W., Witham, P. J., Milanowski, C. A., and Thorn, C. A. *The other 69 percent*.(2009). Center for Educator Compensation Reform. Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education.

Schoeneberger, J. A. and Pulliam, C. L. (2011). *Evaluation of the strategic staffing initiative in Charlotte-Mecklenburg schools.* Center for Research and Evaluation, Office of Accountability, Charlotte-Mecklenburg Schools. Retrieved from http://www.cms.k12.nc.us/cmsdepartments/accountability/cfre/Documents/Stategic Staffing Evaluation Report January 2011.pdf

Raudenbush, S. W., Jean, M. (2012). *How should educators interpret value-added scores?* Knowledge Brief. Carnegie Knowledge Network. Retrieved from http://www.carnegieknowledgenetwork.org/wp-content/uploads/2012/10/CKN_2012-10_Raudenbush.pdf

Roth, P. L., Huffcutt, A. I., and Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology, 88*(4), 694–706.TNTP. (2012). *Keeping irreplacables in D.C. Public Schools: Lessons in smart retention.* Brooklyn, NY: Author. Retrieved from http://tntp.org/assets/documents/TNTP_DCIrreplaceables_2012.pdf

Travers, J. & Christiansen, B. (2010). Strategic staffing for successful schools: Breaking the cycle of failure in Charlotte-Mecklenburg Schools. The Aspen Institute. Retrieved from http://www.aspeninstitute.org/publications/strategic-staffing-successful-schools-breaking-cycle-failure-charlotte-mecklenburg-scho

United States Department of Education, National Center for Education Statistics. (2011). *Beginning teacher longitudinal study (BTLS),* First through third wave preliminary data file (2007–08, 2008–09, 2009–10). Retrieved from http://nces.ed.gov/pubs2011/2011318.pdf

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* The New Teacher Project. Retrieved from http://widgeteffect.org/downloads/TheWidgetEffect.pdf

Zeyu, X., Ozek, U., & Corritore, M. (2012). *Portability of teacher effectiveness across school settings*. National Center for Analysis of Longitudinal Data in Education Research. Retrieved from http://www.caldercenter.org/publications/upload/wp77.pdf

ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.

LOCATIONS

Domestic

Washington, D.C.

Atlanta, GA

Baltimore, MD

Chapel Hill, NC

Chicago, IL

Columbus, OH

Frederick, MD

Honolulu, HI

Indianapolis, IN

Naperville, IL

New York, NY

Portland, OR

Sacramento, CA

San Mateo, CA

Silver Spring, MD

Waltham, MA

International

Egypt

Honduras

Ivory Coast

Kenya

Liberia

Malawi

Pakistan

South Africa

Zambia