Indiana Teacher Appraisal and Support System

Ensuring that teacher evaluation is equitable, effective & efficient

AN ANALYSIS OF INDIANA DISTRICT EVALUATION PLANS

Submitted by:

Dr. Sandi Cole, Ed.D.
Dr. Hardy Murphy, Ph.D.

INTASS
Center on Education and Lifelong Learning
Indiana University

July 6, 2016

This research was supported by grants from the Indiana State Board of Education and the Joyce Foundation.

INTRODUCTION

In 2012, Indiana implemented changes in teacher evaluation required by Indiana Senate Enrolled Act 001 (Public Law 90). A litany of questions concerning best practices in the evaluation of teachers has emerged as districts have responded to this law with newly developed teacher evaluation plans. Among these are questions related to plan development and implementation quality and process. What are the characteristics of high quality evaluation plan development and implementation and what objective measures are districts using to assess student learning, are but two of many questions being discussed in the professional literature concerning teacher evaluation. Adequately answering these questions and others is essential to transforming teacher evaluation into a professional growth experience that develops teachers and ensures student learning.

In 2015, the State Board of Education contracted with INTASS to continue the work of the priorities in the SBOE's Strategic Planning Committee that directly relate to Goal 2 indicators in the SBOE's Strategic Plan. This audit represents one of several deliverables included in that contract.

The purpose of the INTASS analysis provided in this report is to 1) determine the objective measures identified in Indiana school district teacher evaluation plans to assess student learning, 2) identify and review characteristics, components and features of these plans, and 3) inform the ongoing policy discussion of how to ensure the development and implementation of evaluation plans in Indiana effectively support teacher development and student learning needs.

The findings presented in this report were obtained from a review of teacher evaluation plans in Indiana for the 2014-15 and 2015-16 school years, using the principles and practices of the Indiana Teacher Appraisal System of Supports (INTASS) and incorporates the best practices identified in the literature review. Additionally, this audit builds upon the work of the State Board of Education Design Committee, convened after a TNTP review of Indiana's teacher evaluation model and that made specific recommendations on reviewing and revising the teacher evaluation process in the state (TNTP, 2015, SBOE-SPC, 2015).

LITERATURE REVIEW

Plan Development

Developing and implementing high quality teacher evaluation plans with fidelity is more likely to happen if all involved have a clear understanding that the purpose and expectations of the evaluation process are to support teachers in delivering highly effective instruction. Coburn (2005) noted several factors that influence teachers' response to new policy. Clarity about policy goals and means are important, as unclear expectations can cause anxiety and frustration. Another factor is how practical the policy change is to implement. Finally, an important factor in how teachers respond to new policy centers

around the extent the policy change provides teachers with recognizable means and processes to put the change into practice (instrumentality).

Teacher evaluation policy should reflect the purpose of helping all teachers improve (National Council on Teacher Quality, 2011). The Center for Public Education (2013) reported on the elements of good teacher evaluation systems. They noted that an inclusive design and implementation process was a critical element. Ensuring this collaborative approach requires districts to make cultural as well as structural changes to teacher evaluation systems (Cole, Robinson, Ansaldo, Whiteman, & Spradlin, 2012). Engaging key stakeholders in critical reflection and inquiry ensures that those who must implement a new appraisal system have ownership in the development of a district evaluation plan. Teachers and administrators alike must believe that their district evaluation plan is being created in a collaborative culture. Simply adopting a model without engaging in a collaborative process will make implementation problematic. In addition to an inclusive design and implementation process, the Center for Public Education (2013) also noted four other elements of effective teacher evaluation systems: multiple measures, adequate resources and support, data linking teachers to student performance, and classroom observations.

The Council of Chief State Officers (2016) in a recent report on the principles for teacher support and evaluation systems provides ten recommendations for states and districts:

- Regularly communicate the purpose of teacher support and evaluation.
- Build teacher support and evaluations systems on clearly articulated standards and effective teaching practices.
- Clarify the roles and responsibilities of states, districts and schools with regard to teacher support and evaluation systems.
- Ensure support and evaluation is an ongoing process of providing teachers with frequent, action-oriented feedback connected to professional learning.
- Create structures for teachers to work collaboratively to set goals, create and or select measures, and reflect on progress.
- Build the skills of leaders to effectively implement teacher support and evaluation.
- Engage educators in the development of the support and evaluation systems and in its continuous improvement.
- Use multiple, high quality measures to create a comprehensive view of teaching practice.
- Ensure consistency and accuracy of evaluation data.
- Ensure the system is fair, credible and transparent.

Policy and Governance:

Federal policy such as Race to The Top, with its requirements for grant awards and implementation waivers, has played a significant role in the transformation of educator evaluation. In Indiana, legislation followed by state rulemaking, policy and guidance have used the practice of local control at the school corporation level as the cornerstone in the development and implementation of teacher evaluation systems. This practice creates a question concerning the state's role in the development and implementation of high quality teacher evaluation plans.

In guidance to states, Goe, Holdheide and Miller (2014) recommend that States identify measures and conduct research during and after implementation of teacher evaluation and play an active role in conducting research to ensure that the evaluation model(s) is technically sound and therefore defensible, especially in situations in which teacher evaluation results will be used to make personnel and compensation decisions. The Center for Public Education (2013) suggest local school districts need flexibility in designing and implementing evaluation systems to align with district needs, but they also need strong support and guidance from their state. When discussing the balance between local and state control, they state that a medium approach appears best; one in which the state provides a framework in which certain aspects may be mandated by the state but allow other features to be determined locally. The National Council on Teacher Quality (2011) also identifies a problem with a lack of standardization in plan development and implementation and state that it can be difficult to compare teacher quality across districts. There is good reason for states to support districts to implement strong evaluation systems, and little oversight or guidance can be worrisome.

Critical Plan Characteristics, Features, and Processes

Goe et al (2014) published a practical guide for states in creating effective evaluation systems. They identified the following critical components of evaluation plans:

- Ensure that the purpose is to improve teaching and learning.
- Cultivating a strategic communication plan that secures stakeholder investment, and includes gathering feedback on the evaluation plan.
- Use multiple measures to allow for a more comprehensive view of teachers effectiveness based on a variety of evidence.
- Select measures that are valid and reliable for their intended purpose. Measures that have a higher validity may be used with more confidence.
- Ensure factors related to implementation fidelity (such as number of observations, artifacts etc.) are clearly articulated.
- Invest in training of evaluators.
- Ensure data integrity—(Clean data, verifying data, system to collect data, transparency of data).
- Use teacher evaluation data for professional development at the individual, school and district levels.
- Develop factors to determine if the system is effective and implemented with fidelity.

Although evaluation systems have many different components and features, implicit in the list above are several important characteristics that are critical to the successful development and implementation of teacher evaluation plans. Fundamental to the effectiveness of plan development and implementation is the quality of the professional dialogue that is the heart of the evaluation experience. In plan development, this dialogue may be thought of as collaboration. To ensure that the evaluation experience is meaningful and worthwhile evaluators should receive ongoing training (Goe et al, 2014). In the evaluation experience it is the provision of high quality feedback from evaluators that is necessary for teachers to benefit from the evaluation experience. Providing actionable

feedback to teachers, creating professional learning communities where teachers share goals and share responsibility for student outcomes, and forging a system where teachers have opportunity for ongoing professional development can enhance instructional quality. This high quality professional discourse creates in the evaluation process an atmosphere of collegiality in an instructional partnership for teacher and student success. (Hallinger, Heck, Murphy, 2014).

Research Based Teacher Evaluation Rubrics

A number of highly respected teacher evaluation rubrics have emerged over the past several years and are being used by districts in the evaluation process, e.g., Danielson, Marzano, etc. In general these rubrics focus on three or four components of the teaching and learning process (planning, instructional delivery, classroom management and learning environment, and professional development and collegiality). In one form or another many of these rubrics have been a part of the educational environment for a number of years. However, although these rubrics receive widespread acceptance, they cannot overcome the limitations of limited observation time in making a judgement of teaching quality. Unannounced classroom visits (10-15 minutes) by the same evaluator, sampling multiple aspects of the teachers work followed by a face-to-face feedback session is a recommended practice (Marshall, K., 2012). In this way the ratings of teaching quality are supported by multiple instances of observing and noting the instructional practices evidenced through the teacher evaluation rubric.

Objective Measures of Student Learning

The inclusion of student learning in the teacher evaluation experience in a quantifiable manner to inform teacher evaluation ratings is a relatively new development. In 2015, 43 states required objective measures of student achievement be included in teacher evaluations and 17 states required that student growth be preponderant criterion in teacher evaluation (National Council on Teacher Quality, 2015).

The term "objective measure" is used in a variety of situations that range from the technical to the complex and for a variety of research and professional purposes. Objective measurement operates within the research traditions of fundamental measurement theory, item response theory, and latent trait theory. Objective measurement can be achieved and maintained employing a wide variety of approaches and methods. (The Program Committee of the Institute for Objective Measurement, December 2000).

Indiana Code 20-28-11.5-4 requires school corporations to include objective measures of student achievement and growth to significantly inform the final evaluation rating. State Board rule (2012) states the use and weighting of objective measures of student achievement and growth measures shall directly relate to the assessments that most accurately measure student learning according to the following priority:

> (1) Where a mandatory state assessment exists, a school corporation must use it as a measure of student learning. If that state assessment provides individual growth model data (IGM), the district must use it as that teacher's primary measure of student learning.

(2) Where a state assessment does not exist, an assessment developed or procured by the district that is used for common grades or subjects shall be used as a measure of student learning.

(3) Only when there is no state, district or school assessment shall a district utilize class-specific, teacher-created assessments as a measure of student learning for evaluation purposes.

(4) Districts may use multiple student learning measures. If districts choose to use multiple sources of data, the primary measure will carry the most weight in relation to the other student learning measures.

The Indiana State Board of Education Strategic Design Committee (2015), in a review of current experiences in the development and implementation of Indiana teacher evaluation plans also defined objective measures of student achievement and growth. The Committee defined it in the following two statements.

- Objective measures of student learning is an outcome statement of student performance requiring measureable data to support instructional goals.
- Objective measures of student learning allow a valid and reliable assessment of skill and knowledge, attitudes, and opinion with an agreed upon standard or criteria recognized by a properly qualified and/or trained individual or by an individual who is informed in its administration, scoring and interpretation.

The challenge and debate have been focused on whether assessments can, with validity and reliability, measure the impact a teacher has on student learning. Validity is the overarching concept that defines quality in educational measurement. It is the extent to which an assessment measures what it is intended to measure and provide sound evidence (Herman, Heritage, & Goldschmidt, 2011). However, the use of even the most reliable and valid measures is received with some resistance simply because no measure is perfectly reliable or valid. There is always the chance that performance on a test may not represent with certainty the learning that has occurred during the teaching process. The use of multiple measures is important and is meant to compensate for the imperfections of each individual measure and produce more accurate and helpful evaluations (Goe, L., Holdheide, L, & Miller, T. 2014).

Herman, Heritage, & Goldschmidt (2011) establish the basic argument that assessments should be used to measure student growth as a part of teacher evaluation. However, they argue that little attention has been devoted to the *quality* of the student assessments, which is fundamental to the trustworthiness of any teacher value added measure, and state that carefully designed and validated assessments are needed in order to provide trustworthy evidence of teacher quality. They further note that when states understand the requirements that assessments need to satisfy and the essential design features, they can provide needed guidance to districts on quality assessments that should be used in teacher evaluation.

Because the majority of teachers teach in grades and subject areas not covered by typical statewide assessments, districts have often been left to come up with differing ways to

measure student growth for these teachers, even though the statistical properties of these alternatives remain unknown (Coburn 2005). Designing measures of student growth for non-tested areas is an important challenge for states (National Council on Teacher Quality, 2011).

Assessment is critical for all teachers of all subjects and all grades. Non-tested grades and subjects, those not part of a state's accountability system, should not be an afterthought. It is important to determine student learning across all grades and content areas in order for districts and schools ensure a successful educational experience for students and to provide support for all teachers. Good assessment measures regardless of whether they are standardized, "off the shelf", norm referenced, criterion referenced, locally developed, or teacher developed can inform student progress and teacher performance (National Council on Teacher Quality, 2011).

This difference between the availability of standardized measures across content areas is often the source of discontent in the use of student growth in the evaluation process. It is a dilemma that impacts weights chosen for student growth in teacher evaluations across content areas, roles, and assignments. While it is not at all unreasonable that states may make adjustments to component weights as systems mature, treating teachers differently does not seem to be a recipe for teacher satisfaction and trust in the results. Explicit policies should exist for non-tested areas to ensure that there are not lower standards for this group of teachers. States must use caution when using school-wide measures of growth in individual teacher evaluations, as these cannot be a substitute for individual measures (National Council on Teacher Quality, 2013).

The Indiana teacher evaluation model, RISE 2.0, recommended that teachers develop Student Learning Objectives (SLO) to determine student learning. The SLO process is a method of setting measurable goals, or objectives based on the students taught, the subject matter taught, the baseline performance of the students, and the measurable gain in student performance during the course of instruction. State guidance on choosing an appropriate assessment should ensure that the assessment be rigorous, be aligned with state standards, allow comparability across classrooms, and be valid and reliable (Lacireno-Paquet, Morgan, & Mello, (2014). Among districts that use student-learning objectives, the most frequently reported benefit was increased collaboration, whereas value-added models of student growth were perceived as fairer than student learning objectives. Evidence is limited on the reliability and validity of student learning objectives used in early-adopting districts (MuCullough, English, Angus, & Gill. 2015).

In a survey conducted by Murphy, Cole, Pike, Ansaldo, & Robinson, (2014) teachers, principals and superintendents in Indiana believe that teacher effectiveness affects student achievement and that student achievement and growth can be validly measured. However, teachers in this study did not believe that their local plans effectively capture the relationship between effective teaching and student learning.

OBJECTIVE MEASURES OF STUDENT ACHIEVEMENT AND GROWTH USED IN INDIANA TEACHER EVALUATION PLANS

METHODS

The district evaluation plans that were reviewed were submitted to the IDOE in September of 2014. The 2014 plans were selected because it was felt that after two years of implementation the number of districts submitting plans would give a representative response for review. Additionally, at the time of the contract award and initiation of implementation, the 2015 plans had not been submitted.

The INTASS team reviewed 283 plans that were submitted to the IDOE through Legal Standard 12. Ten districts were still under an unexpired contract and therefore, not obligated to implement the requirements of IC 20-28-11.5 and did not submit plan data. Twenty-eight districts submitted plans but had missing student learning and teacher evaluation rubric weight data or the data was unclear. For reporting purposes we are using the number 245 to represent the number of submitted plans for with which all applicable data were clearly present in the plan including specific percentages for the student learning, i.e., achievement and growth, and teacher evaluation rubric and weights.

Below is the process and timelines used for the plan review:

- November 2015: Spreadsheet developed for data entry of objective measures, including:
    - Weights (for 1-4 classes of teachers, depending on district)
        - Observation (TER) rubric
        - IGM data
        - SLO or other standard growth factor
        - School-wide learning (i.e. A-F grade)
    - Indication of name/type of objective measures as specified in plans
- November-December 2015
    - 4 INTASS staffers reviewed a portion of district plans
    - Spreadsheet was populated for weights and objective measures
    - Data merged into single sheet
- December 2015
    - INTASS staffer reviews all plans for name/type of objective measures
- March-April 2016
    - 3 INTASS staffers review weights of all plans
    - Final formulas applied and review completed.

It should be noted here that in addition to an audit of objective measures and their use in the weighting of student learning for rating teachers in the evaluation process, (i.e., growth and achievement), the documentation of the teacher evaluation rubric used and the assigned weight in the rating process offer useful data for documenting plan characteristics.

RESULTS

Objective Measures of Student Achievement and Growth Used in Indiana Teacher Evaluation Plans

There were three distinct types of objective measures in student growth and achievement in the evaluation plans reviewed; IGM, A-F school wide letter grade and SLOs. In the review of the 245 evaluation plans, all plans (100 percent) included a teacher effectiveness rubric. Of the 245 plans reviewed, 212 (87 percent) included the A-F letter grade as an objective measure of student achievement and growth in the evaluation plan and 199 districts (81 percent) included IGM as an objective measure of student achievement and growth in the evaluation plan. Of the 245 plans reviewed, 175 (71 percent) included SLOs as an objective measure of student achievement and growth in the evaluation plan.

The following tables represent analytical findings of the 245 district evaluation plans submitted and reviewed.

Table 1: Represents the number of evaluations reviewed and the number of districts that included the use of a teacher evaluation rubric, IGM and A-F letter grade.

	N=	% (N/245)
Number of Districts Indicating Use of Teacher Effectiveness Rubric (TER)	245	100%
Number of Districts Indicating Use of IGM	199	81%
Number of Districts Indicating Use of A-F letter Grade	212	87%
Number of Districts indicating Use of SLO's	175	71%

Table 2: Represents the highest and lowest weights used for the teacher effectiveness rubric, IGM, A-F letter grade and SLO in calculating summative ratings. The weighted percentage for the use of the Teacher Effectiveness Rubric (not student achievement and growth measures) varies greatly among the district plans. The weighted percentages used for the Teacher Effectiveness Rubric for the 2014-15 school year range from as low as 40 percent to as high as 100 percent. The weighted percentage used for IGM varies greatly among the district plans from a high of 50 percent to a low of 2 percent. The highest percentage used among the districts for A-F was 50 percent and the lowest was 2 percent. For SLO's the highest percentage used among the districts was 60 percent to a low of 2 percent.

Measure	Highest Weights Used	Lowest Weights Used
Teacher Effectiveness Rubric	100%	40%
IGM	50%	2%
A-F School Letter Grade	50%	2%
Student Learning Objectives (SLO)	60%	2%

Most Commonly Used Assessments Cited in District Evaluation Plans:
(Name and Number of Districts)

ECA (116)
Locally Developed Assessments/Final Exam (96)
ISTAR (58 Districts)
DIBELS/mCLASS (57)
NWEA (54)
iMAST (44)
IRead (42)
Acuity (37)
WIDA/LAS Links (32)
Advanced Placement Test Results (24)
Star (26)
Student Portfolio (15)
IEP's/504 Plans (9)
AIMSweb (8)
RIT (7)
WIDA (7)
PSAT/SAT (6)
STI (6)
Lexile (4)
ACT (3)
Grad Rate (2)
Industry Certification Assessments (2)
TerraNova (2)
Student Stakeholder Feedback (2)
Project Lead the Way (2)
INCA test (2)
Aleks (2)
Achieve 3000 (2)
Scholastic Reading Counts (2)
TRC (2)
Core Professionalism (2)
Readistep (1)
Short Cycle Assessments (1)
Harcourt End of Year Reading Assessment (1)
CTE (1)
IKAT (1)
Benchmark Assessments/Running Records (1)
Accu-sess (1)
Easy CBM (1)
Next (1)
CRT (1)
MYP Criterion (1)
ULS Benchmarks (1)
President Challenge Fitness (1)
Dual Credit Percentage (1)
Fitness for Life (1)
School Bullying Assessment (1)
Parental Involvement Data (1)
Accuplacer (1)

Of the 245 plans reviewed, 33 districts are only using one measure of student achievement and growth. IDOE guidance and state law requires multiple measures of student achievement and growth. Nearly one in five (18.7 percent) districts are not using IGM as a measure of student achievement and growth. Of the 33 districts only using one measure of student achievement and growth, 21 are using the school A-F letter grade as the only measure of student achievement and growth. State Board Rule states that districts must use IGM as the primary measure of student learning for all teachers that provide instruction in grades 4-8 English/language arts and/or math. Also of note, only 58 districts indicated that they were using ISTAR, the state alternate assessment for teachers who teach students on alternate assessment. IDOE Guidance states teachers should use the assessment with the most confidence in terms of validity and reliability, which would be ISTAR as it is a state assessment

During the 2014-15 school year, the IDOE conducted onsite monitoring visits of district's implementation of evaluation plans per Indiana's ESEA Flexibility Waiver. During the visits, the IDOE did find districts that were not including IGM as the primary measure of student achievement and growth and issued an individual report to the district. The districts were responsible to provide next steps to the IDOE to ensure IGM would be included in the future. Due to the new federal law, Every Student Succeeds Act (ESSA), the monitoring of district teacher evaluation plans are no longer required.

SLO's were mentioned as part of the student learning weight for all three groups of teachers; however, districts often did not report or describe the measure being used

as a part of the SLO. This was particularly true for those districts using RISE 2.0 or Modified RISE 2.0.

INDIANA SCHOOL DISTRICT ASSESSMENT OF TEACHER EVALUATION PLAN CHARACTERISTICS

METHODS

In March 2016, the Indiana University Center on Education and Lifelong Leaning (IU CELL) contracted with Haley Consulting Services, LLC (HCS), a research and evaluation firm specializing in the review and analysis of programs and services, to assist INTASS in completing an assessment of Indiana school district Teacher Evaluation Plan characteristics. As described in the following narrative, working together with HCS through the fall and spring of the 2015-16 school year, a comprehensive review and assessment of Indiana teacher evaluation plans was accomplished.

In order to address reliability and validity in the review process, project participants developed an assessment tool derived from *The INTASS Educator Evaluation Plan Rubric (2013)*. The INTASS rubric was developed "to assist school districts in developing and implementing high quality educator evaluation systems" (Murphy & Cole 2013). Because the rubric is essentially a tool for *planning*, it was adapted to *enable a reliable document review for the determination/assessment of plan characteristics*. The resulting tool distills the research on developing effective teacher evaluation plans down to 36 key components that the ideal plan should contain. These components all fall under the following categories:

- Intent and Philosophy/Belief Statements (1 component)
- Strategic Communications (2 components)
- Process for Classroom Observations (17 components)
- Weights of measures (2 components)
- Clear timelines/Protocols (1 component)
- System for Measuring Student Learning (4 components)
- System for Collecting, Reporting and Storing Data (3 components)
- Converting Measures Scores to Summative Teacher Ratings (1 component)
- Oversight Process (2 components)
- Professional Development (2 components)
- Forms (1 component)

IU CELL personnel downloaded 2015-16 evaluation plans from the Indiana Department of Education's website (http://www.doe.in.gov/evaluations/evaluation-plans) and converted them into readable files using Adobe Acrobat. A total of 275 district plans were downloaded, of

which 271 were assessed. One plan contained only the cover sheet and no attached plan, one plan was corrupted and unable to open, one plan pertained to only a single school within a district (which is otherwise included in the assessment), and one district submitted their School Improvement Plan, which did not contain the required evaluation plan.

Assessing the plans

Using an Excel spreadsheet to track district scores, the presence or absence of each component in every district plan was recorded using standard dichotomous notation of ones and zeros. Notations were recorded in the spreadsheet to note where in the plans one can find evidence in support of the given score. Finally, key passages were highlighted in the plans to further support the assigned scores. *Appendix A* contains a copy of the codebook used to assess the plans.

Inter-rater reliability

After the first 30 plans were assessed, INTASS personnel conducted an inter-rater reliability test on five randomly selected plans. The result of this test identified areas of inconsistency in the assessment process, which led to a strengthening/clarifications of definitions for several key components. After an additional fifty plans were assessed, another check was conducted on five randomly selected plans. This check revealed that the clarifications increased the reliability of the tool. A final check on 10 randomly selected plans was conducted in order to confirm the inter-rater reliability of the process.

RESULTS

Distribution of scores

Out of a total possible 36 points, scores ranged from two to 32. No district scored a perfect 36 points. The average score was 19, as was the median score. The most common score (the mode) was 18. Using a standard deviation of 5, the distribution of scores was divided into three categories of high, medium and low. Medium scores include those within one standard deviation in either direction from the mean (14 to 24 points); low scores include all those below one standard deviation from the mean, and high scores include all scores above one standard deviation from the mean. Though very slightly skewed to the left, Figure 1 shows a relatively normal distribution of scores for the 271 assessed district plans.

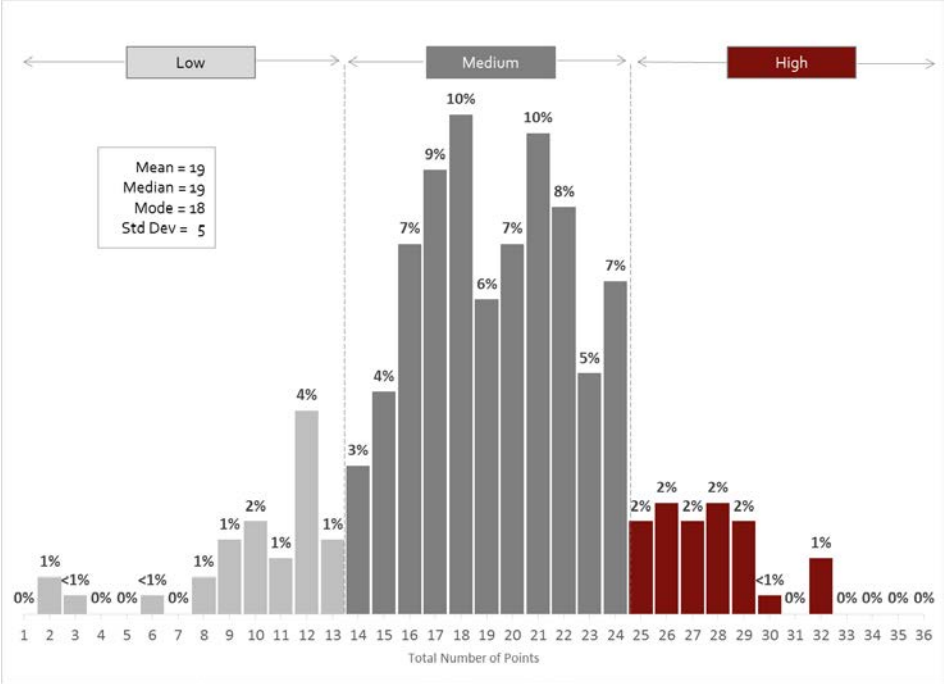

Figure 1. Distribution of scores for district plans. Thirty-one districts scored in the high range, 207 in the medium range, and 33 in the low range.

Five of the six districts that received SBOE recognition (four of which were INTASS districts) scored in the high range, and one scored just below. Five of the six districts that INTASS worked with, independent of the SBOE recognition process, scored in the high range, with one scoring in the medium range (Figure 2).

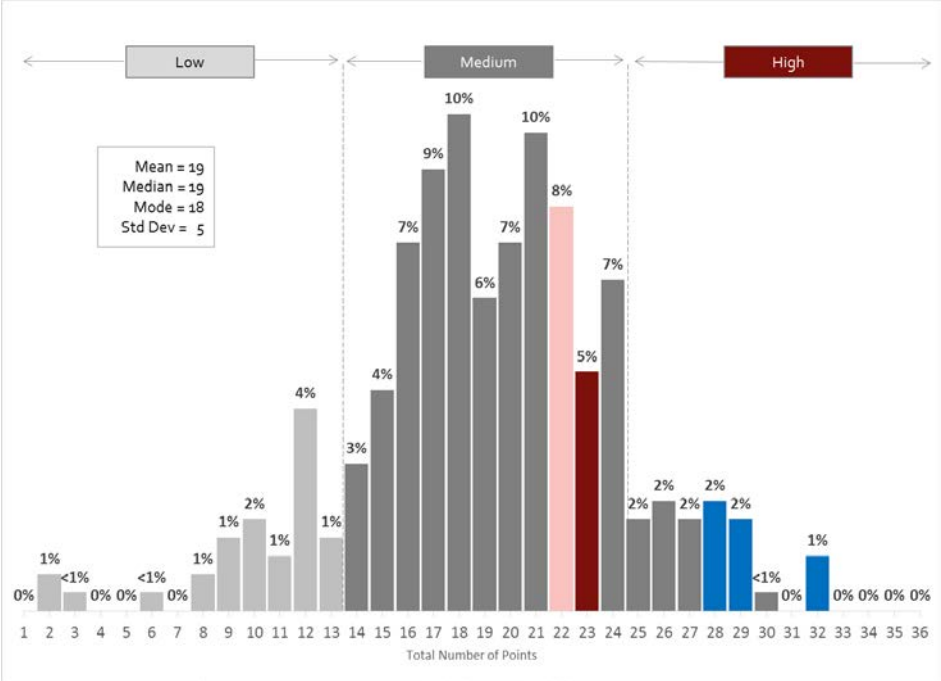

Figure 2. All SBOE recognized districts and INTASS districts scored above average. Pink denotes the score of one SBOE district, crimson denotes one INTASS district, and blue denotes districts that were both SBOE and INTASS

Comparing High, Medium, and Low Scoring Districts

One way to understand the differences between high, medium and low scoring Districts is to look at the percentage of districts within those categories that included each component in their plans. The following charts group similar components and show the percentage of districts that included the corresponding component in their evaluation plans (black bars). To the right of the aggregate black bars, each evaluation plan component is broken down into scores for each of the scoring categories—high, medium or low. To create the breakout bars, each district was assigned to a category based on the total number of points they accumulated through the assessment. From there, totals were calculated for each category and then presented as a percentage of their respective group. For example, in the first chart below, 71 percent of all 271 districts included a purpose and belief statements in their plans, while 97 percent of the 31 high scoring districts included this component, and 77% of the 207 medium scoring districts did the same. Within each grouping, components are sorted from high to low based on the aggregate percentages. This sometimes changes the numbered ordered of the components. The component numbers here correspond to the component numbers in the original INTASS rubric for developing an effective teacher evaluation plan and hold no significance other than to identify them.

Intent/Philosophy/Belief Statements, and Strategic Communications Plan (Figure 3)

In the aggregate, all three of these components could be improved, but it is the gathering of stakeholder feedback on evaluation plans (component 2.2, bottom bars) that stands out the most with 32 percent of all districts including it in their plans. Among the disaggregated scores we see an immediate and strong difference between high scoring districts and low scoring districts. Very few low scoring districts include purpose and belief statements (12%)in their plans, nor do they include information on how they will gather feedback from their stakeholders on the evaluation plan (15%), and nearly half as many low scoring districts specify any strategies for communicating their evaluation plans as do high scoring districts (42% versus 90%).

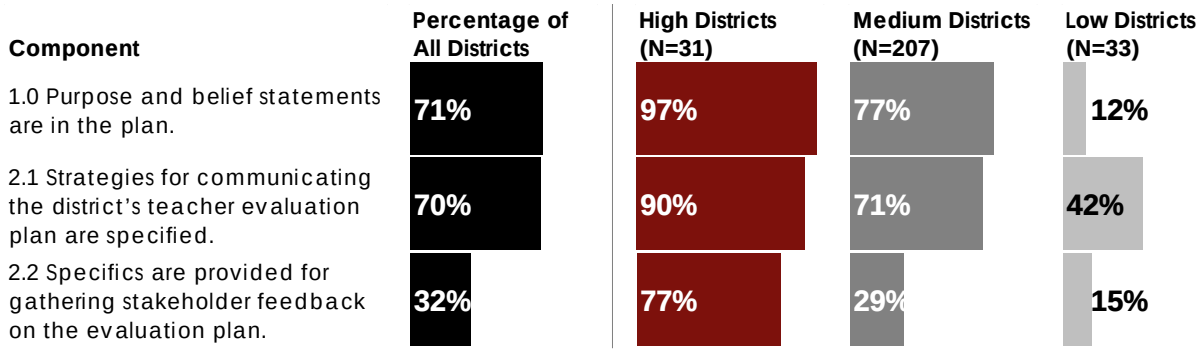

Figure 3. Components related to p**hilosophy/belief statements and communications plans.**

High Quality Teacher Evaluation Rubric (Figure 4)

The teacher evaluation rubric components provided the highest overall averages for all districts combined, as well as for each category—high, medium and low. The lowest overall component for this group was 4.1b. Rubric is applicable for multiple roles and teaching assignments with adaptations. However, even that component was included in the majority of low scoring districts' plans (76%).

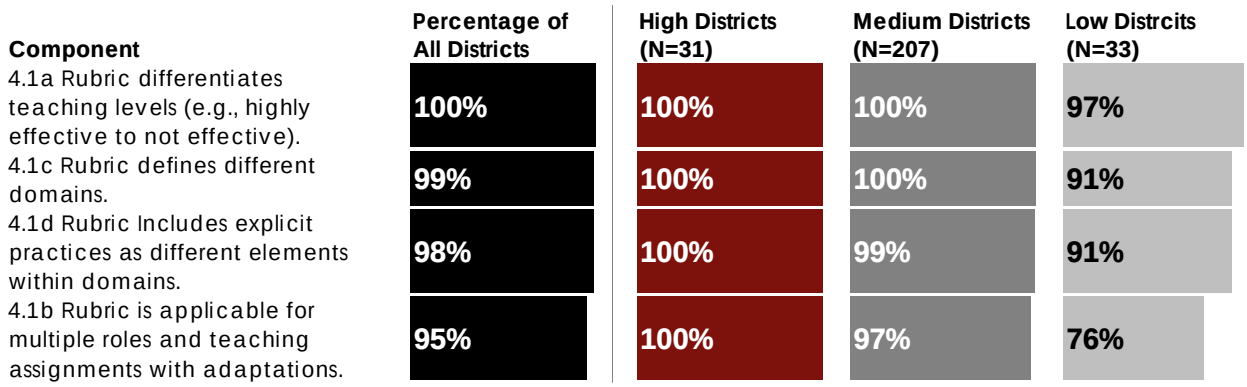

Figure 4. Components related to a h**igh quality teacher evaluation rubric.**

Evaluators and Evaluator Training (Figure 5)

Again, there is a great difference between high scoring and low scoring districts, especially as it relates to clearly defining evaluator roles and responsibilities (component 4.2b, second row). While all of the high scoring districts and nearly the entire medium scoring districts (95%) include this component in their plans, less than one-fifth of the low scoring districts do so (15%). When it comes to training evaluators (4.6a and 4.6b, third and fourth rows), high scoring districts outscore both the medium and low scoring districts, though there is still great room for improvement for them all. Districts were awarded a point for 4.6a if the indicated third party training for their evaluators in their plan. This was done to account for inconsistencies in the language used among districts. It is interesting to note that more low scoring districts include third party training for their evaluators than providing them with a clear understanding of their roles and responsibilities (27% versus 15% respectively). One possibility for the very low scores for 4.6b—yearly renewal training for evaluators—is that many plans used generic language stating that "training was ongoing," which we did not accept as annual renewal.

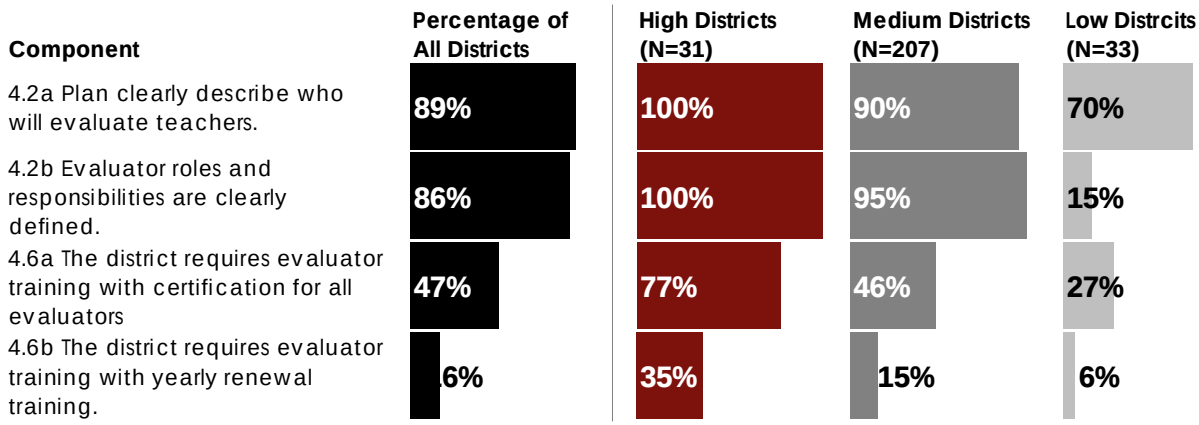

Figure 5. Components related to evaluators and evaluator training.

Observation Timeline and Procedures (Figure 6)

Looking strictly at the aggregate numbers, beyond providing teachers with the number of annual observations that are required (91% of all districts), improvement is needed in terms of clarifying the overall observation process. The description of the process is an area where the high scoring districts excel in all except describing how observations and conferences will be scheduled. Most districts simply stated that these things would happen. While most schools did provide time parameters for providing observation feedback (4.3d, fourth row), the majority did not describe the purpose of the feedback, just that it would be provided. Again, there is a stark difference between high and low scoring districts on this component. Except for describing how conferences will be scheduled, the overall description of the observation process is strength for high scoring districts and a weakness for low scoring ones.

Report continues on next page.

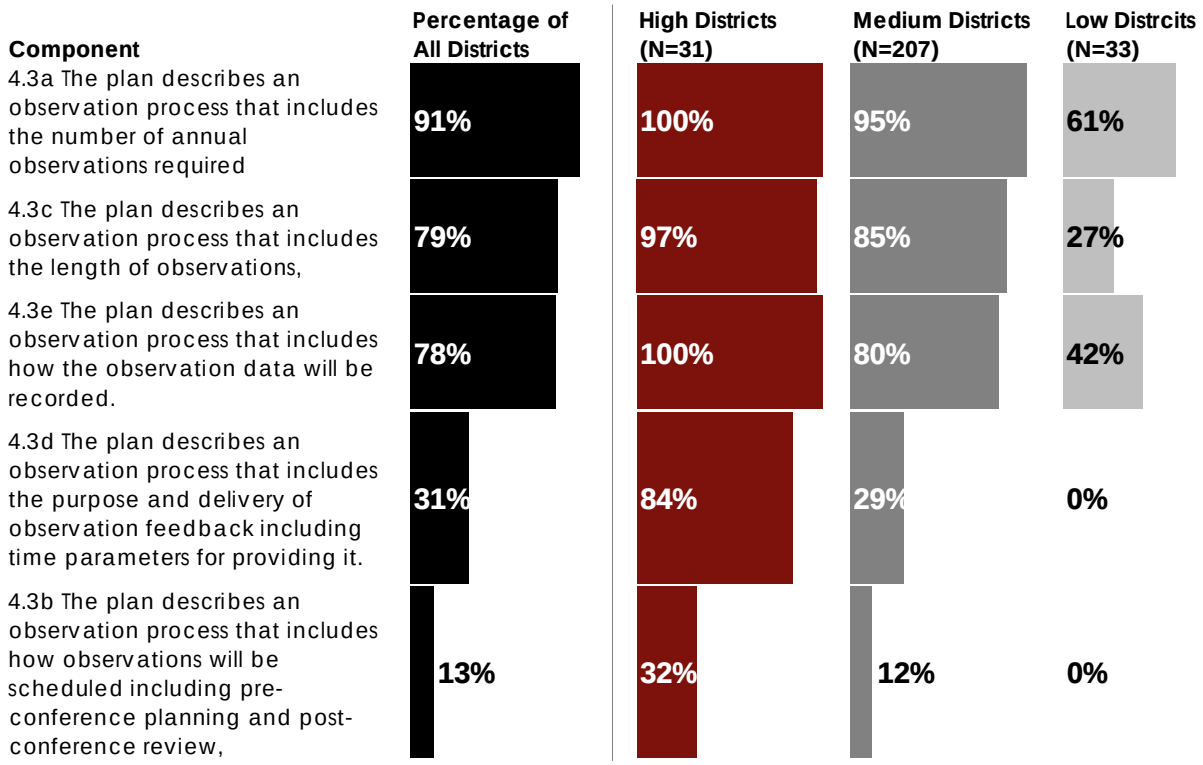

Figure 6. Components related to observation timelines and procedures.

Evidence/Artifacts (Figure 7)

The collection of evidence beyond observations is a component that sets high scoring districts apart from all others. While the aggregate scores in this section are relatively low, all but three (90%) high scoring districts included the collection of artifacts in their plans, and only one low scoring district (3%) did. Medium scoring districts were split down the middle, with 51 percent referencing the practice. Beyond the mention of collecting artifacts, all districts could improve their plans by including how the artifacts will be used (for example, in support of a specific domain or practice in their rubric), and even more so by clearly defining the criteria for collecting the artifacts. Defining criteria would help avoid the random collection of any and all artifacts that a teacher might produce throughout the year.

Report continues on next page.

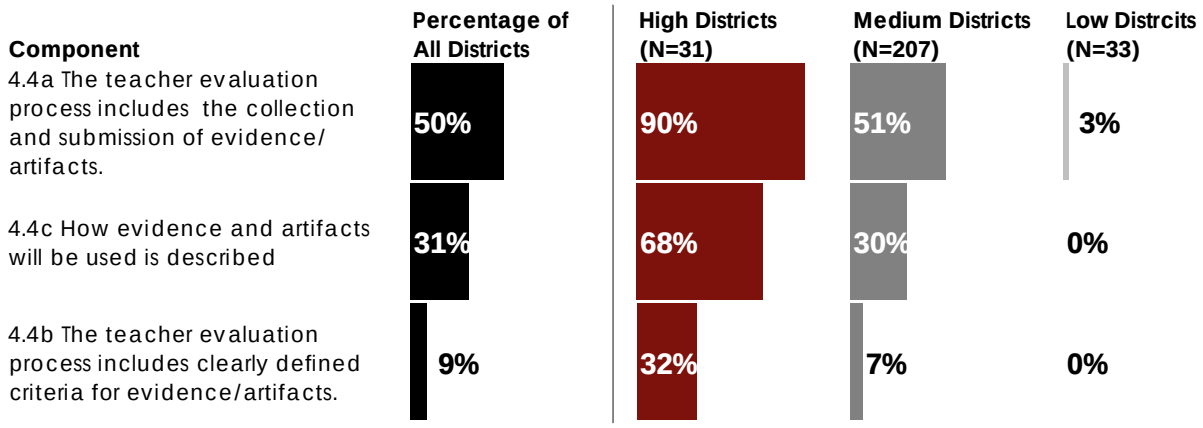

Figure 7. Components related to evidence/artifacts.

Conferences and Meaningful dialogue/feedback (Figure 8)

Including language in plans related to conferences and meaningful feedback was a stumbling block for the vast majority of Indiana school districts. Less than a quarter of all districts require both pre- and post-observation conferences, and less than a fifth note self-reflection as a part of the process. While we do see big differences between high and low scoring districts, these scores are on the low side for the otherwise high scoring districts.

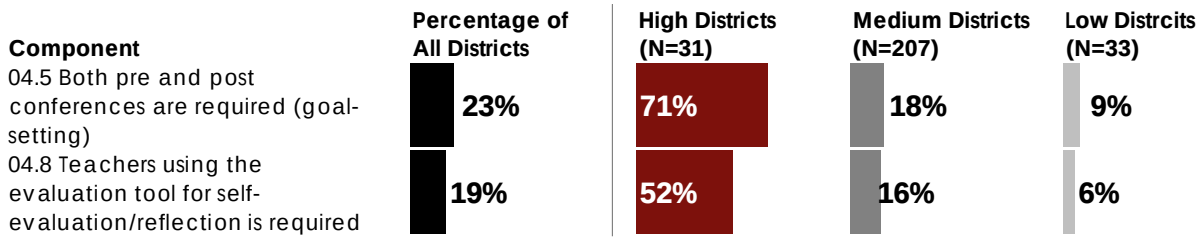

Figure 8. Components related to conferences and meaningful feedback.

Weights of Measures and Summative Scores (Figure 9)

This is a strong group of components for the aggregate of all districts, and is second only to the components concerning an effective observation rubric seen above. Ten districts neglected to include student learning data as part of their summative scores (5.0b in the first row), some of which submitted incomplete evaluation plans that included only the observation rubric. While there is not much difference between high and medium scoring districts, the low districts lag relatively far behind in all but component 5.0b, including student learning data in their summative ratings (which is still 20 percentage points below medium scoring districts). The biggest difference is seen in the third row of data, where the vast majority of high scoring districts (90%) indicate in their plan that they weight all

teachers the same when it comes to student learning data while just over a third of the low scoring districts (36%) note this practice. Many of the districts, from all three categories, that did not receive credit for this used the RISE 2.0 Handbook, or a modified version of it, which differentiated between different types of teachers.

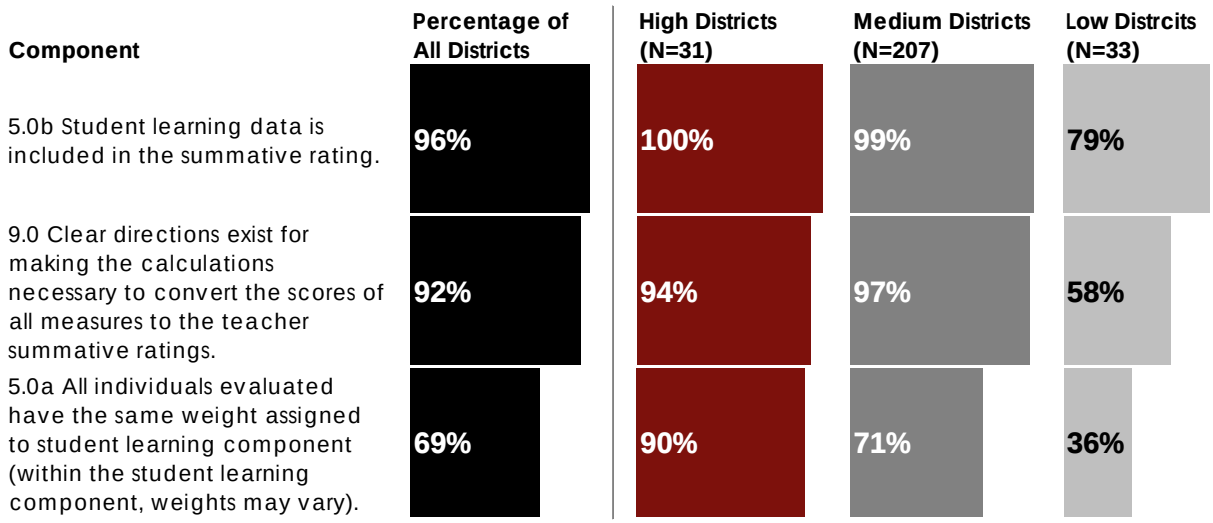

Figure 9. Components related to weights of measures and summative scores.

Measuring Student Learning (Figure 10)

In Figure 10, one percent (equivalent to three districts), state that they include student feedback as part of their student learning scores. Districts are either unaware of this practice altogether, do not agree with including student feedback, or they do not understand the benefits of including it in their process. Either way, it is clear that some effort directly related to implementing this practice is needed if these scores are to increase in the future. The rest of the data in this section is also underwhelming, though quite as drastic. Nearly a third of all districts indicate that they use either published data (usually from the state) or locally produced data to measure student learning, but they do not use a combination of both. As seen in Figure 9 above, there were a few districts that do not indicate that they use student learning data at all; those districts necessarily scored zeros for all Figure 10 components. Selection criteria (7.2 and 7.3, rows two and three) are in great need of clarification by more than three quarters of all districts. While the difference between high scoring and low scoring districts is great, high scoring districts could still use a great deal of improvement in these areas.

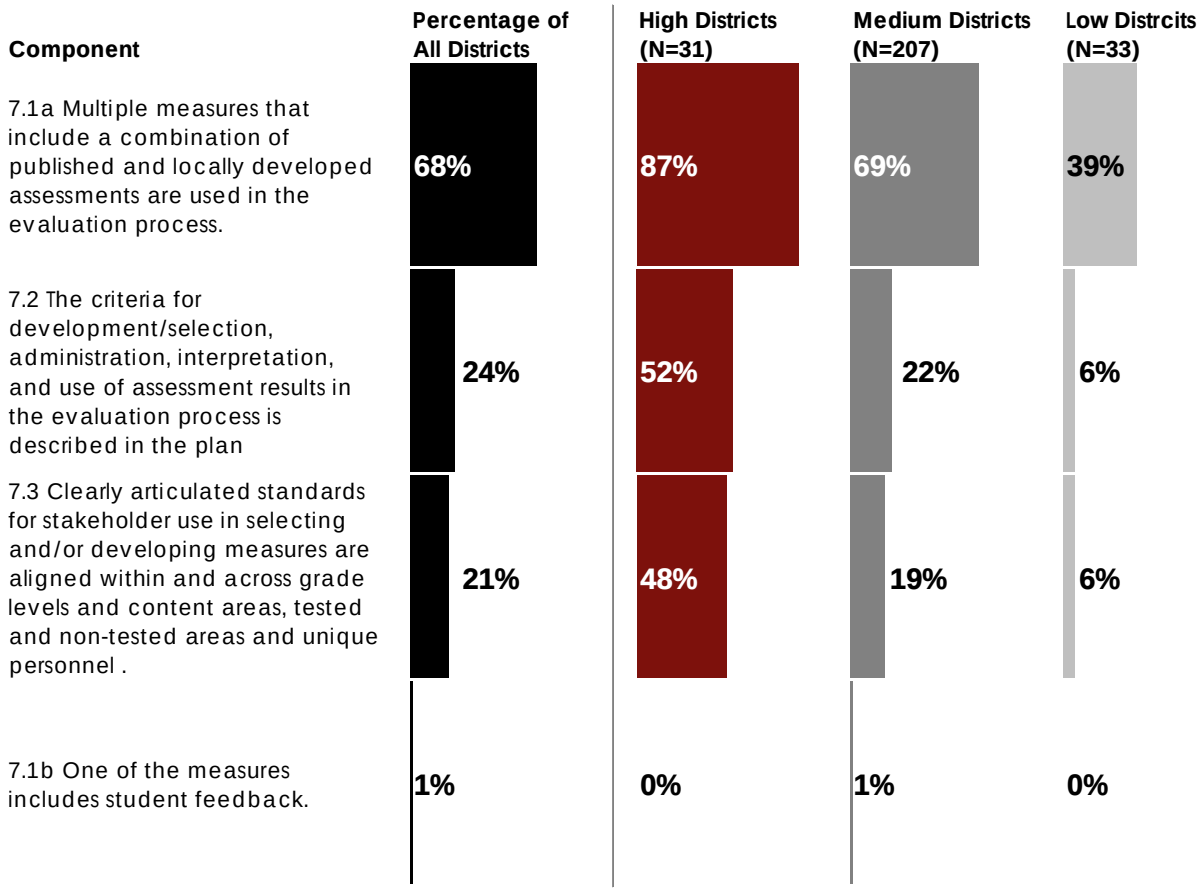

Figure 10. Components related to measuring student learning.

System for Collecting, Reporting and Storing Data (Figure 11)

Practices around data are also a weakness of most districts. While the majority of high scoring districts (77%) identify a data management infrastructure, such as Pivot of Standard for Success, the majority of all districts do not (53%). Very few districts across the board describe guidelines for maintaining the security of testing data (8.4a and 8.4b, bottom two rows).

Report continues on next page.

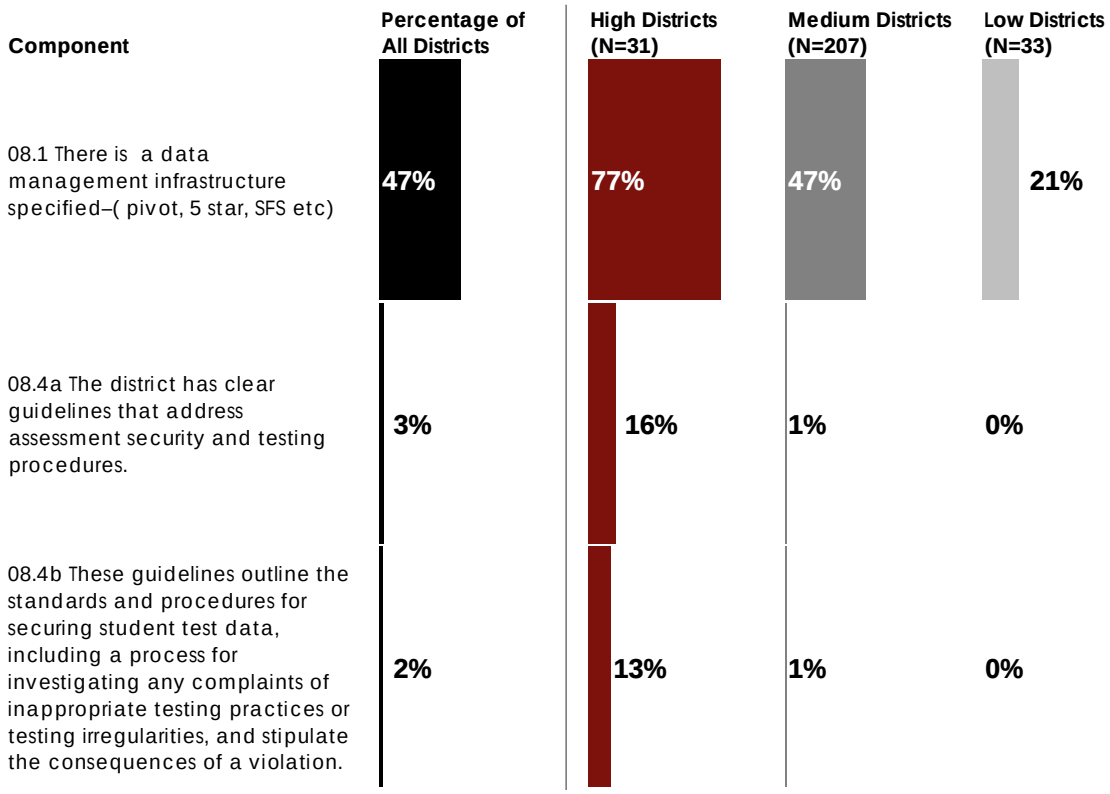

Figure 11. Components related to collecting, reporting and storing data.

Oversight Process (Figure 12)

Evaluation plan oversight is yet another weakness of most district plans. Just over a quarter of the districts indicate that they have a process to monitor and oversee their evaluation plan (28%), and less than a quarter of them (22%) indicate that they include a team that meets regularly to discuss ongoing implementation.

Report continues on next page.

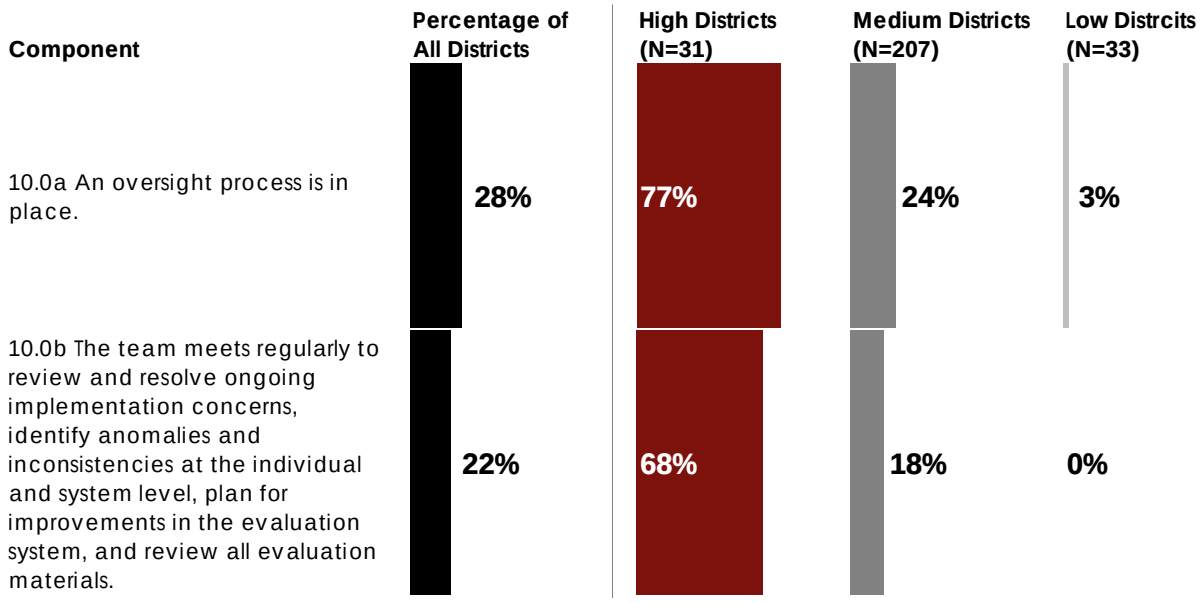

Figure 12. Components related to oversight.

Professional Development (Figure 13)

The difference between the professional development (PD) components below is the difference between PD for an individual teacher (component 11.0a, top row) and PD for groups of teachers—grade level, school level, district level (component 11.0b, bottom row). On the whole, districts did a better job describing clear and specific plans for individual teachers (plans of assistance) than they did describing the relationship between their evaluations and group-level PD. But in both cases, there is great need of improvement. It should be noted that in order to score a point for 11.0a, the plan needed to go beyond state code, which does provide some details (for example, the duration of the plans), but which leaves most of the specifics up to individual districts to define. So, while a district may have included the language from the statute in their plan, they did nothing more to clarify what they do specifically to implement the statute.

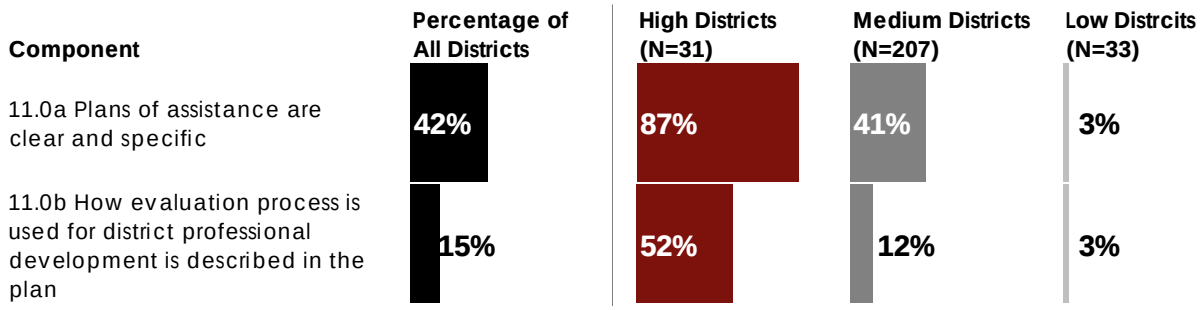

Figure 13. Components related to professional development.

Timelines, Protocols and Forms (Figure 14)

While appearing somewhat unrelated, these two components support the structure of an effective evaluation plan by providing specific timelines that go beyond simply stating that things will happen, and by providing forms that aid in documenting the process. All high scoring districts included or referred to forms in their evaluation plans, while just under two-thirds of the low districts did. More striking, all but two high scoring districts (94%) provided specific timelines, including the months that particular events would take place, in addition to describing clear procedures, while none of the low districts did so.

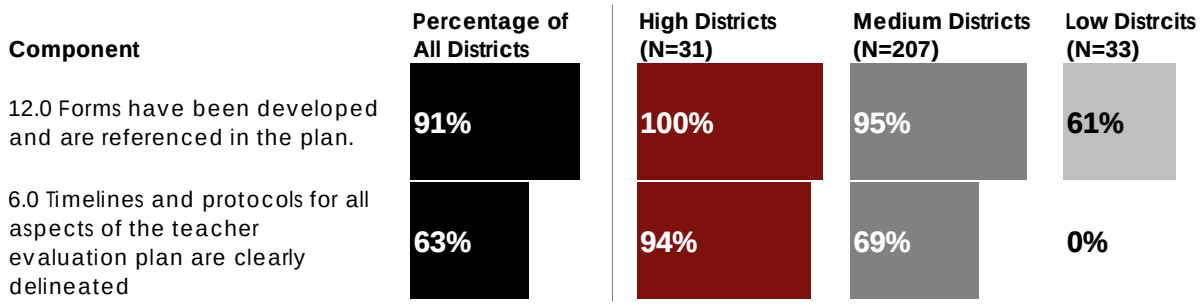

Figure 14. Components related to timelines, protocols and forms.

In the aggregate, district evaluation plans exhibited great strength in five key areas, meaning that more than 90 percent of all districts included the components in their plans:

- Components related to high quality rubrics,
- Describing the number of required observations,
- Including student data in their ratings,
- Clearly describing calculations for summative ratings, and
- Providing forms for documenting evaluations.

In all other areas, district plans show big weaknesses, using 70 percent as the cutoff. This may appear severe, but the literature supports having all of these components as part of an ideal effective evaluation plan (National Council on Teacher Quality, 2013; Goe, L., Holdheide, L, & Miller, T., 2014; Council of Chief State Officers, 2016). Therefore, rigorous standards are called for.

Trying to discover what sets high scoring districts apart from low scoring districts is difficult because we found a difference greater than 30 percentage point in all areas but one (high quality rubrics). Table 1 below provides a summary of strengths, weaknesses and differences that we found in figures above.

Table 1. *Aggregate district plan strengths and weaknesses; and big differences found between high and low scoring districts.*

Figure Number	Aggregate: Strengths (>90%)	Aggregate: Weaknesses (<70%)	High & low districts: Differences (>30%)
3		**Component 2.2**, gathering feedback on the plan from stakeholders	**All components** – purpose/belief statements & communications
4	**All components** related to high quality rubrics		
5		**Components 4.6a** & **4.6b**, certified/third-party trained evaluators and annual evaluation renewal	**Components 4.2b** & **4.6a**, defined evaluator roles/responsibilities and certified/third-party evaluation training
6	**Component 4.3a**, plan includes number of observations	**Components 4.3d** and **4.3b**, how observations & conferences will be scheduled and the purpose of them.	**All components** related to observation timelines and procedures
7		**All components** related to evidence/artifacts	**All components** related to evidence/artifacts
8		**All components** related to conferences and meaningful feedback	**All components** related to conferences and meaningful feedback
9	**Components 5.0b** and **9.0**, student learning data included and clear directions for calculating summative rating	**Component 5.0a**, everyone has the same weighting assigned	**Components 9.0** and **5.0a**, clear directions for calculating summative scores and everyone has the same weighting assigned
10		**All components** related to measuring student learning	**All components** except including student feedback
11		**All components** related to collecting, reporting and storing data	**Component 8.1**, specified data management infrastructure
12		**All components** related to oversight	**All components** related to oversight

Figure Number	Aggregate: Strengths (>90%)	Aggregate: Weaknesses (<70%)	High & low districts: Differences (>30%)
13		**All components** related to professional development	**All components** related to professional development
14	**Component 12.0**, developed forms	**Component 6.0**, clear timelines and protocols	**All components** related to timelines, protocols and forms

An Analysis of Plans to State Code/State Board Rule and IDOE Monitoring Document

An additional analysis was conducted to compare the INTASS assessment tool to State Code/State Board Rule and to the IDOE onsite-monitoring document in relation to the district's evaluation plans. In addition to Component 3.0 of the original *INTASS Educator Evaluation Plan Rubric*, which states that, "All legislative requirements are a part of the evaluation system" (p. 5), eight (8) components of the INTASS assessment tool were identified that correlate directly with State Code/State Board Rule and 11 components that correlate with the IDOE onsite monitoring document. Of the eight components that correlate to State Code/State Board Rule, district plans averaged an 80 percent compliance rate (6.4 out of 8). Forty-three (43) district plans—16 percent—complied fully with all eight components of State Code/State Board Rule. Of the 11 components that correlate to the IDOE onsite monitoring tool, district plans averaged 63% compliance.

Overall, the INTASS Rubric is the most rigorous measure of high quality evaluation plans. In addition to the components included in the State Code/State Board Rule and the IDOE monitoring document, the INTASS Rubric has additional components that address the following areas: 1) the purpose of teacher evaluation; 2) how the plan will be communicated and how the district will gather feedback on the plan; 3) the fidelity of implementation; and 4) data infrastructure, data integrity and data security. The INTASS rubric provides more comprehensive, detailed language related to the expectations for high quality plans.

CONCLUSIONS

Four years after the passage of Senate Bill 1, teacher evaluations in the state of Indiana have changed significantly. Research based teacher effectiveness rubrics are used consistently for teacher evaluation in schools across the state. Additionally, the use of student learning outcomes in the evaluation process is happening in nearly all of the school corporations in the state. However, the quality of teacher evaluation experiences differs throughout the state because of inconsistent plan development

and implementation processes. Different interpretations of educator evaluation requirements expressed in ambiguous language in legislation and code explains some of this inconsistency. Monitoring and oversight with neither incentive nor consequence is also a contributing factor. Plan development and implementation in the state exhibits an erratic profile across a continuum of research based best practices. With the expiration of Indiana's ESEA Flexibility Waiver, the IDOE will no longer take an active role in monitoring teacher evaluation plan implementation across the state, which could result in an even greater inconsistency and non-compliance across the state.

As Indiana moves forward into the era of the Every Child Succeeds Act (ESSA), the findings of this report only heighten the sense of urgency for providing supportive and meaningful evaluation experiences for teachers, the educational professionals having the most impact upon student learning. Ensuring the development and implementation of effective teacher evaluation can be a building block for effective schooling throughout the state of Indiana. Supportive and meaningful monitoring of plan development is critical to this process.

A district's teacher evaluation plan does not guarantee implementation with fidelity. However, unclear expectations and experiences not consistent with what is expected can cause anxiety and frustration (Coburn 2005). Teachers have a right to have certainty around the details of their district's evaluation plan, confidence that what is stated is implemented with fidelity, and the expectation that their evaluation experience will support their growth for the benefit of student learning.

At the policy level, states must be confident that educator evaluations are technically sound and therefore defensible, especially in situations in which teacher evaluation results will be used to make personnel and compensation decisions. The balance between local and state control, where the state provides a framework in which certain aspects are mandated by the state while other features could be determined locally, may be a rout to the ownership of plan requirements at the local level (The Center for Public Education, 2013). However, clear standards and guidelines can inform plan development and implementation and insure that consistency is not sacrificed with allowing this latitude for local responsibility.

One problem with a lack of standardization in plan development and implementation is that it can be difficult to compare teacher quality across districts. In order to ensure that teacher evaluation compliments the accountability and effective teaching and learning, states should develop factors to determine if the system is effective and implemented with fidelity (Goe, Holdheide and Miller, 2014; National Council on Teacher Quality 2011). One way to do this is to ensure consistency and accuracy of evaluation data across the state (The Council of Chief State Officers 2016).

Further, providing assistance to districts to support the organizational culture and climate in the plan development and implementation process will result in improved implementation fidelity in the teacher evaluation process and aligns with the intended purpose of Indiana's legislation—to support teachers and students for success.

ACKNOWLEDGEMENTS

Deep appreciation to the following individuals who supported all aspects of the INTASS District Evaluation Plan Review:

Jim Ansaldo, Ph. D.	CELL/INTASS
Lori Brown	CELL/INTASS
Patricia Campbell	CELL/INTASS
Jean Haley	Haley Consulting Services LLC
Sarah Pies	CELL/INTASS
Mallory Rickbell	CELL/INTASS
James Robinson	CELL/INTASS
Christina Wray	Indiana Institute

REFERENCES

Cole, C., Robinson, J., Ansaldo, J., Whiteman, R., & Spradlin, T. (2012). *Overhauling Indiana Teacher Evaluation Systems: Planning and implementation issues of school districts.* Bloomington, IN: Center for Evaluation and Education Policy, Indiana University

Goe, L., Holdheide, L, & Miller, T. (2014). Practical Guide to Designing Comprehensive Teacher Evaluation Systems. Retrieved from the Center on Great Teachers and Leaders website: http://www.gtlcenter.org/sites/default/files/docs/practicalGuideEvalSystems.pdf

Hallinger, P., Heck, R.H., Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. Educational Assessment, Evaluation and Accountability, 26, 5-28. Doi:10.1007/s11092-013-9175-5

Herman, J.L., Heritage, M., & Goldschmidt, P. (2011). Developing and selecting assessments of student growth for use in teacher evaluation systems. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards and Student Testing. (CREESST)

Indiana State Board Rule (2012). Retrieved from http://www.in.gov/legislative/iac/iac_title?iact=511&iaca=10

Indiana State Board of Education (2015) *Stakeholder Design Committee Recommendations.* Retrieved from http://www.in.gov/sboe/files/SBOE_Stakeholder_Design_Committee_Recommendations.pdf

Murphy, H. & Cole, S. (2013). *INTASS Teacher Evaluation Plan Rubric.* Center on Education and Lifelong Learning: Indiana Institute, Indiana University

Lacireno-Paquet, N., Morgan, C., & Mello, D. (2014). How states use student learning objectives in teacher evaluation systems: A review of state websites (REL 2014-013). Washington, D.C.: U.S. Department of Education, Institute of Education Sciences,

Marshall, K. (2012). Fine-tuning teacher evaluation. Educational Leadership 70(3), 50-53.

McCullough, M., English, B., Angus, M.H., & Gill, B. (2015). Alternative student growth measures for teacher evaluation: Implementation experiences of early-adopting districts (REL2015-093). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and

Regional Assistance, Regional Education Laboratory Mid Atlantic. Retrieved from http://ies.ed.gov/ncee/edlabs.

Murphy, H., Cole, C., Pike, G., Ansaldo, J., & Robinson, J. (2014) Indiana *Teacher Evaluation: At the crossroads of implantation.* Bloomington, IN: Center on Education and Lifelong Learning, Indiana University.

Principles for Teacher Support and Evaluation Systems. Retrieved from the Council of Chief State Officers website: http://www.ccsso.org/Documents/2016/Principles%20for%20Teacher%20Support%20and%20Evaluation%20Systems.pdf

State of the States 2011: Trends and Early Lessons on Teacher Evaluation and Effectiveness Policies. Retrieved from the National Council on Teacher Quality website: http://www.nctq.org/dmsStage/State_of_the_States_Teacher_Evaluation_and_Effectiveness_Policies_NCTQ_Report

State of the States 2013: Connect the Dots-Using evaluations of teacher effectiveness to inform policy and practice. Retrieved from the National Council on Teacher Quality: http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report

State of the States 2015: Evaluating Teaching, Leading and Learning. Retrieved from the National Council on Teacher Quality website: http://www.nctq.org/dmsView/StateofStates2015

TNTP (2015). *Final Recommendations for Changes to Indiana's Teacher Evaluation System.* Retrieved from http://www.in.gov/sboe/files/TNTP_Board.pdf

Trends in Teacher Evaluation: How states are measuring teacher performance. (2013). Retrieved from the Center for Public Education website: http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/Trends-in-Teacher-Evaluation-At-A-Glance/Trends-in-Teacher-Evaluation-Full-Report-PDF.pdf

National Center for Education Evaluation and Regional Assistance, Regional Laboratory Northeast & Islands. Retrieved from http://ies.ed.gov/ncee/edlabs

APPENDIX A – TEACHER EVALUATION PLAN ASSESSMENT CODE BOOK

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
1.0 Intent and Philosophy/Belief Statements	1.0 Intent and Philosophy/Belief Statements	01.0 Purpose and belief statements are in the plan.	
2.0 Strategic Communication Plan	*2.1 Communication structures*	*02.1 Strategies for communicating the district's teacher evaluation plan are specified .*	*Yes – mention on school board agenda/minutes; communication with teachers that they're having an evaluation*
	2.2 Process for stakeholder feedback	*02.2 Specifics are provided for feedback - (big picture feedback, not individual feedback)*	*Yes – if there's an oversight committee; example: they do a survey every year to get feedback on the plan*
4.0 Process for Classroom Observations	*4.1 High Quality Teacher Evaluation Rubric*	*04.1a Differentiates teaching levels (e.g., highly effective to not effective)*	*Must be explicit*
		04.1b Rubric is applicable for multiple roles and teaching assignments with adaptations.	*Yes – multiple roles/rubrics present; RISE, TER, Marzanno, Danielson, McCrel, TAP*
		04.1c Defines different domains	
		04.1d Includes explicit practices as different elements within domains.	

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
4.0 Process for Classroom Observations	*4.2 Evaluators*	*04.2a Plan clearly describes who will evaluate teachers.*	
		04.2bEvaluator roles and responsibilities are clearly defined.	*Yes – if it's clear from the whole document what evaluators' role is*
	4.3 Observation Timeline and Procedures	*04.3a The plan describes an observation process that includes the number of annual observations required*	
		04.3b The plan describes an observation process that includes how observations will be scheduled including pre-conference planning and post-conference review,	*Yes – explicitly notes HOW they'll schedule, not just that there will be an observation (or that it will be scheduled)*
		04.3c The plan describes an observation process that includes the length of observations,	*Yes – must be explicit* *No – "short" or "long"*
		04.3d The plan describes an observation process that includes the purpose and delivery of observation feedback including time parameters for providing it, and	*Yes – needs to say something about WHY they're meeting/giving feedback AND the timeframe for giving it (within X days)*
		04.3e The plan describes an observation process that includes how the observation data will be recorded.	*Yes – must include something about how/where – forms or computer software*

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
4.0 Process for Classroom Observations	*4.4 Evidence/Artifacts*	*04.4a The teacher evaluation process includes the collection and submission of evidence/artifacts.*	*Yes – collecting artifacts is noted but not the procedures for collecting them*
		04.4b The teacher evaluation process includes clearly defined criteria for evidence/artifacts.	*Yes – is it purposeful*
		04.4c How evidence and artifacts will be used is described	*Yes – how they're going to score it*
	4.5 Pre and Post Conferences	*04.5 Both pre and post conferences are required (goal-setting)*	*Yes – baseline conference/beginning of year conference if it's used for goal-setting; No – end of year conference does not count as a post-conference*
	4.6 Evaluator Training	*04.6a The district requires evaluator training with certification for all evaluators*	*Yes – certification or third-party training noted*
		04.6b The district requires evaluator training with yearly renewal training.	*Yes – any kind of ANNUAL training*
	4.8 Meaningful dialogue and feedback	*04.8 Teachers using the evaluation tool for self-evaluation/reflection is required*	*Yes – PD plan works if there's evidence of self-reflection and required by all*

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
5.0 Weights of Measures	5.0 Weights of Measures	05.0a All individuals evaluated have the same weight assigned to student learning component (within the student learning component, weights may vary	
		05.0b Student learning data is included in the summative rating	*Yes – any indication of student data (not just teacher observation)*
6.0 Clear timelines/Protocols	6.0 Clear timelines/Protocols	06.0 Timelines and protocols for all aspects of the teacher evaluation plan are clearly delineated	*Yes – looking for months when things happen/number of days; No – only indicating semesters*
7.0 System for Measuring Student Learning	*7.1 Multiple Measures*	*07.1a Multiple measures that include a combination of published and locally developed assessments are used in the evaluation process.*	*Yes – has BOTH published and locally developed*
		07.1b One of the measures includes student feedback.	*Yes – they ask students for feedback*

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
7.0 System for Measuring Student Learning	*7.2 Criteria for Selection and Development of Measures*	*07.2 The criteria for development/selection, administration, interpretation, and use of assessment results in the evaluation process is described in the plan*	*Yes – must meet ALL of the above*
	7.3 Process for Selection and Development of Measures	*07.3 Clearly articulated standards for stakeholder use in selecting and/or developing measures are aligned within and across grade levels and content areas, tested and non-tested areas and unique personnel.*	*No – if no from 7.2* *Yes – especially for locally-developed: all teacher of X doing it the same way*
8.0 System for Collecting, Reporting and Storing Data	*8.1 Infrastructure*	*08.1 There is a data management infrastructure specified– (pivot, 5 star, SFS etc)*	*Yes – unidentified software program*
	8.4 Assessment Security and Procedures	*08.4a The district has clear guidelines that address assessment security and testing procedures.*	*Yes – looking for data security*
		08.4b These guidelines outline the standards and procedures for securing student test data, including a process for investigating any complaints of inappropriate testing practices or testing irregularities, and stipulate the consequences of a violation.	*Yes – looking for process for investigating complaints or breaches*
9.0 Converting Measure Scores to Summative Teacher Ratings	9.0 Converting Measure Scores to Summative Teacher Ratings	09.0 Clear directions exist for making the calculations necessary to convert the scores of all measures to the teacher summative ratings.	*Yes – if they provide percentage breakdowns for each component of summative score*
10.0	10.0 Oversight	10.0a An oversight	*Yes – must be a*

Plan Component	Plan Sub-component	Attribute	Notes/clarifications made after first inter-rater reliability check
Oversight Process	Process	process is in place.	*committee in place; No – superintendent reviews the process/annually reviewed by school board*
		10.0b The team meets regularly to review and resolve ongoing implementation concerns, identify anomalies and inconsistencies at the individual and system level, plan for improvements in the evaluation system, and review all evaluation materials.	*Yes – annually counts*
11.0 Professional Development	11.0 Professional Development	11.0a Plans of assistance are clear and specific	*Yes – more than the "canned" language from code*
		11.0b How evaluation process is used for district professional development is described in the plan	*Yes – looking for something bigger picture, not just how it relates to individual teachers*
12.0 Forms	12.0 Forms	12.0 Forms have been developed and are referenced in the plan.	