

Education Policy Brief

Overhauling Indiana Teacher Evaluation Systems: Examining Planning and Implementation Issues of School Districts

Cassandra M. Cole, James N. Robinson, Jim Ansaldo, Rodney S. Whiteman, and Terry E. Spradlin

VOLUME 10, NUMBER 4, SUMMER 2012

CONTENTS

Introduction	1
Superintendents' Survey on Teacher Evaluation	2
Summary of Survey Results	2
Factors Influencing Evaluation Model Selection	3
Consensus and Communication Among Stakeholders	5
Professional Practice Measures	7
Student Learning Measures	8
Post-Evaluation Features	9
Employing a Process for Implementation	9
Conclusions and Recommendations	10
Policy Perspectives	
Dr. Hardy Murphy	12
Lynn Holdheide	13
Ryan Snoddy	14
Authors	15
Acknowledgements	15
References	15
Web Resources	16

UPCOMING POLICY BRIEFS . . .

- ✓ *The Evolution of Indiana's School Accountability System*
- ✓ *Charter Schools: Charter Revocation Procedures*
- ✓ *Trends in Virtual Learning in the U.S. and Indiana*

INTRODUCTION

As one part of a broad reform package, the 2011 Indiana General Assembly passed Senate Enrolled Act 001 (Public Law 90), a bill that significantly changed the way Indiana teachers will be evaluated and compensated. An earlier Center for Evaluation & Education Policy (CEEP) policy brief, *Revamping the Teacher Evaluation Process* (Whiteman, Shi, & Plucker, 2011), provides an extensive review of this law and the various evaluation models available to school districts to use. School districts throughout Indiana worked extensively during the 2011-12 school year, with guidance from the Indiana Department of Education, to put together evaluation plans that align with the legislation.

The requirement for districts to move from a traditional form of evaluation and compensation that has been the practice in Indiana for decades, to a system that more directly links teacher performance to compensation has been a conundrum for many districts. On the one hand, many believe that the system of teacher evaluation needs to change. On the other hand, the comfort with the old way of doing business and the fear of attaching compensation to performance have caused consternation for many educators. The shift to a new way of thinking about teacher evaluations has required educators to reflect on their core convictions about instruction and to come to a shared understanding of concepts such as "growth," "accurate," "fair," and "effective." For both administrators and teachers alike, the legislation has evoked a multitude of emotions, created many a debate, provided for reflective conversations, and required collaborative leadership among teachers and administrators.

The legislation and subsequent guidance from the state have ensured that district conversations have focused on various aspects of evaluation models. Shall we adopt the state's evaluation model, referred to as RISE, or adapt it? Should we consider using another performance rubric such as Danielson or Marzano? What about adapting the Evanston model? What will the state allow or not allow with student learning measures? These and other questions related to evaluation models are important. However, as most of these conversations have focused on the selection of a model, more consideration should be given to the process by which models, or components of models, are chosen, implemented, and refined once implemented.

This brief is the first of a two-part series that will explore how Senate Enrolled Act 001 is being operationalized across Indiana school districts. In February 2012, a survey was sent to Indiana superintendents to gauge their attitudes and beliefs about the legislation. The findings, which are discussed in this brief, found that superintendents (a) agreed that the evaluation process in Indiana needed to change, (b) believed that student achievement and growth can be measured, (c) agreed that evaluations should be linked to student learning, (d) were less sure that teacher evaluations should be linked to compensation, (e) believed that evaluations should inform professional development, and (f) were concerned about the implementation of the new system/models and needed professional development. This brief also discusses the features of quality evaluation plans and introduces essential elements of a planning process that ensures equitable, effective, and efficient plans for evaluating educators.

The second brief, which will be released in fall 2012, will present challenges and implications of Public Law 90 for professional development, policy, and teacher education programs.

SUPERINTENDENTS' SURVEY ON TEACHER EVALUATION METHOD

Data for this brief were collected via the Center for Evaluation & Education Policy (CEEP) and the Center on Education and Lifelong Learning (CELL) Superintendents' Survey on Teacher Evaluation Issues. This survey instrument was administered through an online format and distributed through the Indiana Association of Public School Superintendents (IAPSS). IAPSS members were notified of the survey by an email sent by Dr. John Ellis, Executive Director of IAPSS. The email informed IAPSS superintendents about the CEEP-CELL partnership in collecting, analyzing, and reporting data for an education policy brief; provided a link and URL to complete the survey; assured confidentiality; and indicated that superintendents could opt for receiving additional information about technical assistance and support for developing and implementing teacher appraisal systems.

The survey instrument was originally open from January 17 to January 31, 2012, but the window was then extended to February 3, 2012. The instrument contained 17 items, 13 of which are summarized in this brief. The remaining four items asked participants about their interest in receiving a copy of the brief or technical assistance, and for participants' contact information in order to receive that information and support. Of all IAPSS members receiving the survey (approximately 291), 205 began taking the survey, and 179 completed it. Some participants could be colleagues within the same school district, as some duplicate school district identifiers were entered into the survey. Of the 179 completed surveys, 173 had unique identifiers for school districts.

Survey items included a variety of response types. Questions 2, 3, 7, and 11 asked for opinions on Likert-type scales. Questions 5, 8, and 13 were binary response questions. Questions 4, 10, 12, and 14 allowed respondents to select multiple responses, while questions 6 and 9 asked respondents to select from a list. Additionally, questions 6, 10, 12, and 14 allowed respondents to select "Other" and explain in an open response. Open responses were coded by CEEP researchers and are included in the results summary.

SUMMARY OF SURVEY RESULTS

A. Attitudes Toward Teacher Evaluation and Development (Q.2a – Q.2k)

Superintendents were asked for their level of agreement with various statements related to evaluation and development on a 7-point Likert scale, anchored by "Strongly Agree" and "Strongly Disagree." Results are summarized in Table 1. The survey found that 98.3% of superintendents strongly agreed or agreed that teacher effectiveness affects student achievement, 74.9% strongly agreed or agreed that student achievement can be validly measured (19.6% somewhat agreed), and 72.1% strongly agreed or agreed that student academic growth can be validly measured (19.6% somewhat agreed). When asked if they believed that teacher evaluation should be linked to student growth, 55.3% of respondents strongly agreed or agreed, 33.5% somewhat agreed, and 8.4% somewhat disagreed or disagreed. A total of 63.1% of respondents strongly agreed or agreed that the relationship between teaching and learning can be accurately applied to an evaluation of teaching, with 29.6% somewhat agreeing with the statement.

Table 1. Attitudes Toward Teacher Evaluation Topics (n=179)

	Strongly Agree	Agree	Somewhat Agree	Neither Agree nor Disagree	Somewhat Disagree	Disagree	Strongly Disagree
Teacher effectiveness affects student achievement	87.7%	10.6%	0.6%	0.0%	0.0%	0.0%	0.6%
Student achievement can be validly measured	24.0%	50.8%	19.6%	1.7%	3.4%	0.0%	0.6%
Student academic growth can be validly measured	17.3%	54.7%	19.6%	3.4%	4.5%	0.0%	0.6%
Teacher evaluation should be linked to student growth	14.5%	40.8%	33.5%	2.8%	3.4%	5.0%	0.0%
Instruction can be accurately evaluated and judged	24.0%	49.7%	22.9%	1.1%	0.0%	2.2%	0.0%
The relationship between teaching and learning can be accurately applied to an evaluation of teaching	14.5%	48.6%	29.6%	2.2%	1.1%	2.2%	1.7%
Teacher evaluation should be tied to merit/compensation	6.7%	21.2%	31.8%	10.1%	7.8%	7.8%	14.5%
Prior to the new law, the teacher evaluation processes in Indiana needed improvement	35.2%	34.1%	17.9%	6.1%	2.2%	1.7%	2.8%
An effective teacher evaluation system informs professional development	36.3%	46.4%	12.3%	3.4%	0.0%	0.0%	1.7%
An effective teacher evaluation system drives professional development	34.6%	44.7%	12.8%	3.9%	1.1%	0.6%	2.2%
Indiana's new law regarding teacher evaluation will result in improved teaching and learning	13.4%	28.5%	22.9%	16.8%	5.6%	7.3%	5.6%

Responding superintendents were more skeptical about linking evaluation to compensation. When asked if they believed that teacher evaluation should be tied to merit or compensation, 27.9% strongly agreed or agreed, 31.8% somewhat agreed, and 22.3% disagree or strongly disagreed.

When asked if the teacher evaluation process in Indiana needed to be improved prior to passing SEA 001, 69.3% of the respondents strongly agreed or agreed. A total of 82.7% strongly agreed or agreed that an effective teacher evaluation system informs professional development, and 79.3% strongly agreed or agreed that an effective teacher evaluation system drives professional development.

Respondents were more cautious about outcomes of Indiana’s new teacher evaluation laws. When asked if Indiana’s new law would result in improved teaching and learning, only 41.9% strongly agreed or agreed, 22.9% somewhat agreed, 16.8% neither agreed or disagreed, 5.6% somewhat disagreed, and 12.8% disagreed or strongly disagreed.

B. Familiarity with Senate Enrolled Act 001-2011 (Q.3 – Q.4)

Survey participants were also asked about their familiarity with SEA 001. Specifically, they were asked about their familiarity with the SEA 001 requirements for teacher evaluation systems and with the options available for developing a plan for teacher evaluation in the respondents’ own school districts. A total of 89.4% responded that they were extremely familiar or adequately familiar with the requirements and options. Participants were then asked how district leadership and staff became familiar with SEA 001 requirements and the resources used to inform the school district’s selection of a teacher evaluation system. Respondents could select any of the options that applied. These responses are summarized in Table 2.

C. Selection of Teacher Evaluation Systems (Q.5 – Q.7)

Superintendents were asked about school district selection of teacher evaluation systems for the 2012-13 school year. Of surveyed superintendents, 80.4% indicated their school districts had selected a teacher evaluation model for the next academic year and 19.6% responded their school districts had not.

If respondents indicated they had selected a teacher evaluation model, they were directed to questions 6 and 7. A total of 143 respondents (n = 143) were directed to questions 6 and 7. In an open response question, participants were asked which teacher evaluation model they selected. A total of 73.4% responded they selected RISE, 4.9% selected the IN-TASS/Evanston-Skokie School District 65 model (IN-TASS), 2.8% selected The System for Teacher and Student Advancement (TAP), and 18.9% indicated they selected another model.

Respondents who selected “Other” for question 6 were given an opportunity to describe which model their school districts selected. These open responses were coded by CEEP researchers. Of these responses, 8.4% of the question 6 total (n = 143) indicated they selected a modified version of the RISE model, 2.1% selected the Danielson Framework, 2.1% selected McREL (one respondent indicated a combination of Danielson and McREL), and 3.5% selected a locally developed model. Of all superintendents responding to question 6, a total of 81.8% indicated that they will use RISE or a modified version of RISE.

FACTORS INFLUENCING EVALUATION MODEL SELECTION

Superintendents whose school districts had selected a teacher evaluation model were asked about the extent to which certain factors may have influenced the selection of those models. Effects were evaluated on a 5-point Likert scale with “Major Effect” and “No Effect” as anchors. All responses are summarized in Table 3.

A total of 39.9% responded that sufficient state and local support for teacher evaluation system development and adoption was a major effect for selecting a model and 42.0% responded it was a moderate effect. Teacher support for adopting a system was a major effect for 30.1% of superintendents and a moderate effect for 44.1%. When considering the influence of sufficient training for implementation, 40.1% responded this was a major effect and 41.5% responded it was a moderate effect (n = 142). Transparency of evaluation systems was a major effect for 33.8% of superintendents, a moderate effect for 38.0%, and was neutral for 21.1% (n = 142).

The response with the greatest variation was Q.5e, which asked how much of an effect the ease of use/flexibility of the system had on the selection of a teacher evaluation model. A total of 26.2% indicated ease of use/flexibility was a major effect, 35.5% indicated it was a moderate effect, 23.4% indicated it was neutral, 7.8% indicated it was a minor effect, and 7.1% indicated it had no effect on selecting a teacher evaluation model (n = 141).

When asked about the effect the cost of the system had on the selection, 38.5% specified it was a major effect, and 35.0% specified a moderate effect. Reliability and relevance of the system to improve student achievement was a major effect for 34.3% of superintendents and a moderate effect for 40.6%. A total of 37.8% responded that reliability and relevance of the system to judge teachers fairly was a major effect on model selection,

Table 2. Sources of Information About Senate Enrolled Act 001-2011 (n=179)

Response	Count	Percentage
Read the legislation	166	92.7%
Attended workshops/seminars	172	96.1%
Participated in webinars	133	74.3%
Held discussions with staff	165	92.2%
Discussed requirements of the law with other superintendent(s)	165	92.2%
Spoken with IDOE officials or reviewed information on the IDOE website	148	82.7%

Table 3. Influences on Teacher Evaluation System Selection (n=179)

	Major Effect	Moderate Effect	Neutral	Minor Effect	No Effect
Sufficient support for teacher evaluation system development and adoption (state and locally)	39.9%	42.0%	14.0%	1.4%	2.8%
Teacher support for adoption of system	30.1%	44.1%	17.5%	3.5%	4.9%
Sufficient training for implementation	40.1%	41.5%	12.0%	3.5%	2.8%
Transparency of system	33.8%	38.0%	21.1%	3.5%	3.5%
Ease of use/flexibility of system	26.2%	35.5%	23.4%	7.8%	7.1%
Cost of the system	38.5%	35.0%	15.4%	6.3%	4.9%
Reliability and relevance of the system to improve student achievement	34.3%	40.6%	20.3%	2.8%	2.1%
Reliability and relevance of the system to judge teachers fairly	37.8%	41.3%	15.4%	2.8%	2.8%
Reliability and relevance to improve teacher effectiveness	39.2%	42.0%	14.7%	2.1%	2.1%

Table 4. Stakeholders Included in Teacher Evaluation Plan Development (n=179)

Response	Count	Percentage
Closed Responses		
Parents	17	9.5%
Students	12	6.7%
Teachers	174	97.2%
Association Leaders	166	92.7%
Principals	179	100.0%
Central Office Staff	173	96.6%
Data Management/IT	96	53.6%
Community	18	10.1%
Other*	18	10.1%
Coded Open Responses**		
School Board Members	12	6.8%
Technology Director	1	0.6%
TAP Leadership at the State Level	1	0.6%
Guidance Counselor	1	0.6%
Consultant	3	1.7%

* Respondents were encouraged to type in an open response if they selected "Other."

** Open responses were coded and listed here.

and 41.3% responded it was a moderate effect. Finally, when asked about the effect that reliability and relevance to improve teacher effectiveness had on selection of a teacher evaluation system, 39.2% indicated it was a major effect, and 42.0% indicated it was a moderate effect.

D. Implementation Timeline (Q.8 – Q.9)

All survey participants (n = 179) were asked about the status of their school districts' collective bargaining agreements (CBA). The status of the CBA required implementation of new teacher evaluation plans by July 1, 2012, for 51.4% of respondents. For those superintendents who indicated they did not have to implement a new plan by July 1,

2012, 65.9% responded their school districts will have an additional year under the CBA, 15.3% responded they will have two more years, and 18.8% responded they will have three or more years.

E. Teacher Evaluation System Development and Implementation (Q.10 – Q.14)

Superintendents were asked which stakeholder groups have been or will be a part of the plan development process. Survey participants could select any applicable stakeholders from the supplied list; they could also select "other" and then write in any additional applicable stakeholders. These open responses were coded by CEEP researchers. Responses are summarized in Table 4.

F. Concerns about Plan Development and Implementation

Survey participants were asked about their concerns with teacher evaluation plan development and implementation. The questions were posed by asking, "What level of concern do you have regarding each of the following items?" on a 5-point Likert scale. All responses are summarized in Table 5.

A total of 67.6% of respondents were extremely concerned or very concerned about resources to conduct classroom observations, 59.2% were extremely concerned or very concerned about resources to collect student performance data, 59.7% were extremely concerned or very concerned about resources to provide training for evaluators, and 60.9% were extremely concerned or very concerned about resources to provide training for staff. Resources for the increased teacher compensation component received the highest levels of concern in question 11, with 69.2% indicating they were extremely concerned and 20.7% indicating very concerned (a total of 89.9%). When asked about the level of concern for the implementation timeline, 52.5% were extremely concerned or very concerned, but 25.7% were only somewhat concerned.

Superintendents were asked about concerns for building the capacity for understanding among school personnel, and 59.8% responded they were extremely concerned or very concerned, but 25.7% responded they were somewhat concerned. Additionally, 45.8% were extremely concerned or very concerned about communication to key stakeholders, with 26.8% somewhat concerned and 20.7% slightly concerned.

The last three items of question 11 refer to school districts' teacher evaluation systems

and continuing support, guidance, and compliance. When asked about levels of concern for ongoing support for professional development, 77.6% responded they were extremely concerned or very concerned. A total of 70.4% responded they were extremely concerned or very concerned about clear guidance concerning the interpretation of the new law, and 18.4% responded they were somewhat concerned. Finally, 54.2% of survey participants were extremely concerned or very concerned about alignment of the new law with policy, 27.9% were somewhat concerned about alignment, and 17.9% were slightly concerned or not at all concerned.

Participants were asked how the requirement for annual teacher evaluation through classroom teacher observation was going to be achieved in their school districts. Respondents could select any applicable response or type in an open response (see Table 6). Of those responses, 24.6% reported that building-level administrators would be responsible for conducting annual teacher classroom observations. A total of 87.2% indicated that school districts will work within existing staff and revise job descriptions or re-classify that existing staff. Only 10.1% reported they plan to hire outside contractors to conduct the required observations, but a total of 15.6% reported they plan to hire additional staff to complete the observations.

G. Teacher Evaluation Data

Management

Survey participants were asked about plans for data management required of a new teacher evaluation system, and 80.4% indicated they have or will implement changes

to data management infrastructure (n = 179). Respondents that indicated they have or will implement data management changes (n = 144) were asked about the changes they have or will implement. Superintendents could select all applicable changes from a list. They also had the option of choosing “Other” and describing the changes. CEEP research assistants coded respondents’ descriptions, and responses are summarized in Table 7.

Of the respondents stating they have or will implement a data management change, 61.8% indicated they will supplement existing data management systems, 54.2% will purchase new hardware or software to manage data, 23.6% will contract with an external firm or foundation, and 11.1% reported they will share data management resources with another school district.

CONSENSUS AND COMMUNICATION AMONG STAKEHOLDERS

In this section, we examine the results of the CEEP-CELL Superintendents’ Survey on Teacher Evaluation Issues by highlighting key themes found in the data. First, we examine issues facing superintendents as they work to develop consensus and communication among stakeholders in their districts during development and implementation of evaluation plans. Next, we reflect on superintendents’ attitudes regarding professional practice measures and the resources implicated therein. This section is followed by a discussion of findings in the domain of student learning measures as an indicator of teacher effectiveness. Lastly, we explore how superintendents envision the outcomes and

repercussions of the new law on teacher development and district planning.

By and large, it appears that the superintendents in Indiana are well aware of Senate Act 001, and that they have enlisted a host of approaches to become informed about the law and its implications. Over 90% of the respondents have read the legislation and attended workshops or seminars about it. A similar percentage of district leaders have engaged in discussions with professionals in their districts, as well as convening with their peers in the superintendency across the state. A total of 82.7% of superintendents have spoken with officials from the Indiana Department of Education to obtain guidance regarding the requirements of the legislation, and 74.3% have participated in webinars.

At the same time, as noted, over 80% of respondents expressed concern about their capacity to build understanding about the new evaluation systems among their school personnel. Close to 60% indicated they are “extremely” or “very” concerned.

The kinds of understanding that may be indicated in the superintendents’ concerns reflect the findings of Newmann, King, and Youngs, (2000), who reference school capacity as containing five elements:

- Teachers’ knowledge, skills, and dispositions;
- Professional community;
- Program coherence;
- Technical resources; and
- Principal leadership.

For Newmann and his colleagues, the central theme is that these elements interactively contribute to greater or lesser school

Table 5. Level of Concern Toward Plan Development and Implementation (n=179)

	Extremely Concerned	Very Concerned	Somewhat Concerned	Slightly Concerned	Not at all Concerned
Resources to conduct classroom observations	33.0%	34.6%	20.7%	7.8%	3.9%
Resources to collect student performance data	25.7%	33.5%	26.8%	10.1%	3.9%
Resources to provide training for evaluators	31.8%	27.9%	21.2%	13.4%	5.6%
Resources to provide training for staff	29.6%	31.3%	23.5%	11.7%	3.9%
Resources for increased compensation component	69.3%	20.7%	5.6%	2.2%	2.2%
Implementation timeline	34.1%	18.4%	25.7%	17.9%	3.9%
Building the capacity for understanding among school personnel	24.6%	35.2%	25.7%	10.6%	3.9%
Communication to key stakeholders	15.6%	30.2%	26.8%	20.7%	6.7%
On-going support for professional development	39.1%	38.5%	10.6%	8.9%	2.8%
Clear guidance concerning the interpretation of new law	40.8%	29.6%	18.4%	7.3%	3.9%
Alignment of new law with policy	26.3%	27.9%	27.9%	11.2%	6.7%

Table 6. Allocation of Human Resources for Plan Implementation (n=179)

Response	Count	Percentage
Closed Responses		
Additional Personnel	27	15.1%
Contracting	18	10.1%
Revising Job Descriptions	98	54.7%
Re-classifying Staff	51	28.5%
Other*	72	40.2%
Coded Open Responses**		
Administrators will Conduct Evaluations	27	15.1%
Principals will Conduct Evaluations	12	6.7%
Not sure	7	3.9%
Rework Current System, Work Within Current System and Staff	7	3.9%
Remain the Same as Current Evaluation Assignments, Practices	3	1.7%
Conduct Evaluations via Master and Mentor Teachers through TAP	1	0.6%
New Hiring Contingent on Funds	1	0.6%
New Training, Equipment, Tools, Software for Staff Evaluators	5	2.8%
Principals and Assistant Principals	5	2.8%
Superintendent Assist in a Primary or Support Function	4	2.2%
Combined Responses***		
Building-level Administrators	44	24.6%
Hiring Additional Personnel	28	15.6%
Revising Job Descriptions and Work within Current System and Staff	105	58.7%
Revising Job Descriptions, Reclassifying Staff, and Work within current system & staff	156	87.2%

* Respondents were encouraged to type in an open response if they selected "Other."

** Open responses were coded and listed here.

*** Combined responses were derived from inferred overlap in closed responses and the coded open responses.

capacity, and that professional development activities (and, seemingly, teacher evaluation systems) should attend to all of them in concert, when possible. Rosenholtz (1989) presents a similar model for what she calls "learning-enriched schools." In such schools, "shared school goals" drive "teacher learning and teacher collaboration," which should influence "teacher commitment and student learning." Rosenholtz found evidence that a sixth element, "teacher certainty," was present among teachers able to express confidence that school goals, collaboration, and professional development were actual factors in both student outcomes and their own commitment to the educational enterprise.

A key component in the success of any school reform effort lies in resolving the tension involved in managing both structural and cultural changes (Fullan, 2001). Fullan argues that federal and state policies tend to represent structural changes, in this case the adoption of accountability measures designed around compliance. Elmore (2004) concurs, and suggests that mere compliance with accountability standards is unlikely to promote lasting change so long as the cultures within schools remain static:

Table 7. Data Management Changes Due To Teacher Evaluation Requirements (n=144)

Response	Count	Percentage
Closed Responses		
Supplement Existing	89	61.8%
Purchase New as Primary	75	52.1%
Contract	33	22.9%
Share with Other Districts	16	11.1%
Other*	15	10.4%
Coded Open Responses**		
Five-Star Technology SDA	2	1.4%
Agreement with Foundation for Data Management	1	0.7%
Creating own Evaluation Software, Programs	2	1.4%
Purchasing Software Packages, Supporting Hardware	3	2.1%
Creating New Staff Position to Manage Data and Data Entry	1	0.7%
Not Sure	6	4.2%
Combined Responses***		
External Data Management	34	23.6%
Purchasing New Hardware or Software	78	54.2%

* Respondents were encouraged to type in an open response if they selected "Other."

** Open responses were coded and listed here.

*** Combined responses were derived from inferred overlap in closed responses and the coded open responses.

...the attitudes, values, and beliefs of individual teachers and administrators...are key factors in determining the solutions that schools construct to the accountability problem. Put bluntly, many educators simply do not believe that they have the capacity to influence student learning in the ways that external accountability systems suggest they should. Hence, external accountability systems will be relatively powerless in the absence of changed conceptions of individual responsibility and collective expectations within schools. (Elmore, 2004, p. 199)

Developing cultures of shared responsibility within schools points toward the need for capacity building among stakeholders (Fullan, 2001), a need the superintendents appear to recognize.

The relationship between and among stakeholders in a school district reflects the micropolitics of the local setting. Malen (1995) discusses these interactions in terms of how power is exchanged among specific stakeholder roles. Of interest here are the dynamics of “professional-patron” transactions, which Malen describes as occurring in formal and informal arenas, and which posit varied strategies for interested actors. Our survey findings indicate that, in the formal arena (i.e., participation on teacher evaluation planning committees), just over 26% of the districts feature participants other than school personnel, such as parents, students, and community representatives. However, it is unclear what distinction, if any, respondents made between community and parents. Regardless, Malen’s meta-analysis paints a less-than robust picture of actual engagement. Rather, participation on such a committee tended to manifest in “cordial, ceremonial exchanges that reflect(ed) and reinforce(d) a traditional pattern of power wherein professionals...control school policy” (Malen, 1995, p. 149-150). Follow-up studies of Indiana superintendents relative to the current teacher evaluation work would do well to explore the comportment of external stakeholders as members of district planning teams.

The National Comprehensive Center for Teacher Quality (Coggshall, 2007; Goe, Holdheide, & Miller, 2011) offers that districts implementing new teacher evaluation systems attend to four key components of a “communication framework” to ensure a shared understanding. This framework advocates for a clear communication plan; articulated goals for the evaluation system; shared definitions of teacher quality and effective-

ness; and specific means to measure teacher effectiveness, qualifications, and expertise. In 2011, the National Council on Teacher Quality (NCTQ) suggested, similarly, that effective evaluation systems feature “comprehensive communication plans” that lend to public understanding of the goals and processes of the mechanisms by which teachers are measured. Two decades earlier, Darling-Hammond, Wise, & Pease (1983) also found that successful operation of teacher evaluation systems required district leaders and teachers to hold common assumptions regarding the purpose and definitions of teacher quality.

Districts face different timelines for the adoption of teacher evaluation systems based upon the status of their collective bargaining agreements with their local teacher associations. Roughly half of responding districts will be required to implement their evaluation plan in the current year, while others have a year or more to prepare their teachers and administrators. At first glance, although those districts with more time might appear to be better positioned to address concerns around capacity, Indiana’s law nonetheless mandates that all districts submit a plan to the Department of Education by July 1, 2012. This may mean that opportunities to communicate effectively around model selection were truncated; further study is warranted, as the survey found that 80.4% of responding superintendents’ districts have adopted a teacher evaluation model. A wide range of factors affecting the decision about model adoption received moderate or major consideration, according to the superintendents. Describing the character and content of these deliberations is, however, beyond the scope of this study.

PROFESSIONAL PRACTICE MEASURES

The need for teacher evaluation systems to rest upon a foundation of multiple measures has been advocated by a host of sources (Coggshall, 2007; Daley & Kim, 2010; Goe, Bell, & Little, 2008). This section addresses and expands on the superintendents’ responses to the selection and implementation of professional practice measures, defined here as indicators of actions and activities of teachers that are not directly observed in student performance outcomes. Examples of such measures include, but are not limited to, planning, classroom management, instructional practices, and attendance to professional responsibilities. This brief also examines the use of student performance measures, in the following section.

First, let us consider the superintendents’ attitudes about the value of evaluating professional practice. An overwhelming majority (98.3%) either agreed or strongly agreed that teacher effectiveness affects student achievement. With perhaps a degree of tepidity, superintendents report that the evaluation process can reveal the relationship between teaching and learning (63.1% agreed or strongly agreed; 29.6% somewhat agreed). Furthermore, more than 96% of the superintendents agreed (in varying degrees) that instruction can be accurately judged. The research on these effects is nascent, but in a study of schools in Cincinnati, Milanowski (2004) found “small-to-moderate positive correlations” between the Danielson-based teacher evaluations and student performance outcomes. Marzano (2011) has also shown positive correlations between his Teacher Evaluation Model and mathematics and reading achievement scores. In any event, the superintendents surveyed seem to believe that evaluating professional practice is both possible and informative to expectations of student outcomes.

Despite having confidence that teacher evaluations can inform the linkage between teacher action and student outcomes, around two thirds of the survey respondents revealed high levels of concern regarding the capacity to conduct the classroom observations necessary to generate said evaluations, in addition to securing the training for personnel responsible for completing them. To get a sense of the task before school leaders, an understanding of the common elements that appear in quality professional practice measures may be helpful.

The first consideration involves the frequency of classroom observations. The state of Indiana’s RISE model (Indiana Department of Education [IDOE], 2011) requires at least one extended (40-minute) and one short (10-minute) observation per year, although more are recommended. The recommendation for multiple observations is far-reaching (Bill and Melinda Gates Foundation, 2012; Daley & Kim, 2010; Danielson, 2007; Marzano, 2011; The New Teacher Project, 2010). These writers acknowledge the time expenditures and staffing necessary to conduct multiple observations, and The New Teacher Project (2011) offers specific guidelines and suggestions for managing resources effectively to meet this need. As indicated earlier, the superintendents share some concern about managing this component of teacher evaluation, and follow-up studies should attempt to describe the various solutions districts devise as they move forward with their plans.

A further commitment of resources lies in the consistent recommendation for pre- and post-observation conferences (Danielson, 2007; Marzano, 2011), although RISE simply suggests beginning and end-of-year conferences between teachers and evaluators (IDOE, 2011).

Another key component of most professional practice measures is a rubric (Danielson, 2007; IDOE, 2011; Marzano, 2011) that describes, defines, and rates types of teacher behaviors. Typically, one finds the array of teacher behaviors arranged into categories or domains, with a fair amount of similarity and/or overlap to be found from model to model. For example, Danielson (2007) features four domains of teacher performance: Planning & Preparation, the Classroom Environment, Instruction, and Professional Relationships. Marzano (2011) suggests that teacher observations look for Classroom Strategies and Behaviors, Planning and Preparing, Reflecting on Teaching, and Collegiality and Professionalism. Indiana's RISE model (IDOE, 2011) offers the Teacher Effectiveness Rubric, featuring three "components" (Planning, Instruction, and Leadership), leaving evaluation of professional responsibilities to a standalone category.

A final consideration in the implementation of professional practice rubrics lies in the need for evaluators to be trained. Indiana Senate Enrolled Act 001 (2011) requires that individuals conducting evaluations receive training, and the literature is clear on this as well. Danielson (2007) considers trust to be a critical characteristic of a teacher evaluation system designed to promote professional learning, and such trust is, in part, fostered by the knowledge that there exists "sufficient training to ensure that administrators can make consistent judgments based on evidence of practice" (Danielson, 2007, p. 177). Building a system based on trust will require resources, as close to 60% of the responding superintendents who expressed higher levels of concern regarding evaluator training appear to recognize.

STUDENT LEARNING MEASURES

This section reexamines survey results related to the characteristics and use of student learning measures in light of educational literature focusing on features of quality systems for teacher evaluation. Survey respondents demonstrated strong agreement with the notion that student achievement and growth can be measured validly. Though they demonstrated strong overall agreement,

respondents were somewhat less confident in the belief that teacher evaluation should be linked to student growth and that the relationship between teaching and learning can be accurately applied to an evaluation of teaching. Similarly, the educational literature offers precautions in the use of student learning measures to determine teacher effectiveness. Survey respondents also demonstrated significant concern regarding availability of resources to support professional development for evaluators and teachers. Again, the literature mirrors this concern.

Indiana is one of 23 states that require teacher evaluations to include measures of student growth and one of 17 states specifying that such measures will "significantly" inform teacher evaluations (NCTQ, 2011). The educational literature suggests that focusing on student growth, rather than achievement, makes evaluation systems more fair to teachers of students who begin the year below grade level (Goe, Holdheide, & Miller, 2011). As stated, survey responses implied that Indiana superintendents are confident that student achievement and student academic growth can be validly measured.

Though they demonstrated strong overall agreement, superintendents were slightly more uncertain in their beliefs that teacher evaluation should be linked to student growth and that the relationship between teaching and learning can be accurately applied to an evaluation of teaching. In other words, although educational leaders feel confident in the field's ability to identify effective instructional practice and measure student learning, connecting the two is a brand new endeavor. The educational literature agrees; there is little evidence regarding implementation and outcomes of teacher evaluation systems (NCTQ, 2011).

Evidence indicates that value-added models (VAMs)—processes that mathematically model student growth and attribute it to particular teachers—can be effective in identifying student growth trends and effective instructional practices (Whiteman et al., 2011). However, use of VAMs as summative tools in teacher evaluation remains controversial. There exists concern over year-to-year fluctuations in teacher ratings (Goe, Holdheide, & Miller, 2011), use of standardized assessments that were not designed for use in VAMs (Sartain, Stoelinga, & Brown, 2011; Whiteman et al., 2011), and reliable integration of these VAMs into existing evaluation systems (Whiteman et al., 2011). The educational literature indicates that VAMs increase in their reliability over time, when several years of data can be aggregated (Curtis & Wiener, 2012).

Several other issues limit the degree to which the field is able to reliably connect instructional practice and measures of student learning. For example, as schools increase their implementation of instructional arrangements whereby students are taught by multiple educators, it becomes increasingly challenging to attribute student growth to particular professionals (Sartain et al., 2011). In addition, there is uncertainty around undermining the original intent of diagnostic and formative assessments by using them for evaluative purposes (Curtis & Wiener, 2012). One of the largest areas of concern is around developing assessments in curricular areas for which state-level assessments do not exist.

Were the controversy over VAMs resolved, the fact would remain that the majority of teachers do not teach a subject or grade level that is currently tested at the state level (Curtis & Wiener, 2012; Goe, Holdheide, & Miller, 2011; NCTQ, 2011; Sartain et al., 2011). Although some states recommend that teachers and evaluators work together to develop student learning objectives (SLOs) for such subjects and grade levels, the process has been criticized as being labor intensive (Curtis & Wiener, 2012). It is likely that development of student learning measures for non-tested areas accounts for at least some of survey respondents' significant concern regarding the availability of resources to collect student performance data.

Despite the field's uncertainty, evidence indicates that instructional practice and student learning are connected. For example, a two-year study of Chicago's Excellence in Teaching program found a strong, positive relationship between classroom observation ratings of teachers and measures of student growth (Sartain et al., 2011). Moreover, it is argued that teachers themselves believe that student growth should be used as an indicator of their performance (Morris, 2012). Rosenthal (1985) argues that teacher effectiveness increases when teachers feel certainty around establishing concrete goals for instruction and knowing when their efforts have produced the desired results around student learning. Thus, measures of student learning play a central role in resolving ambiguity in teacher performance.

A degree of consensus exists, then, around the necessity of incorporating measures of student learning into teacher evaluation systems. Recommendations center on the use of multiple measures as a means of ensuring fairness to teachers (Goe, Holdheide, & Miller, 2011; Morris, 2012). The Indiana Department of Education recommends consideration of several different types of measures:

- Student Learning Measures Linked to Individual Teachers,
- Student Learning Measures Linked to Entire Schools,
- Instructional Practice Measures,
- Professionalism Measures, and
- School and Educational Community Measures.

(IDOE, 2011)

POST-EVALUATION FEATURES

This section reexamines survey results related to post-evaluation features in light of educational literature focusing on features of quality systems for teacher evaluation. Survey respondents demonstrated significant concern regarding the availability of resources to provide training for evaluators and staff. Both respondents and the educational literature agree that the connection between professional development and teacher evaluation is vital in order for the latter to improve teaching and learning. Respondents also demonstrated skepticism and concern regarding the use of teacher evaluation data in making decisions regarding teacher compensation. Also, although the literature debates the merits of evaluation systems that focus either on eliminating or improving ineffective teachers, Indiana superintendents fear that they lack the resources to implement either strategy.

Survey respondents demonstrated significant concerns regarding the availability of resources to provide training for evaluators. The educational literature indicates that when teachers receive quality feedback from classroom observations, both instructional practice and student learning improve (Jerald, 2012). The literature also demonstrates that principals and others evaluators require training to develop skills that will guide teacher reflection, give teachers quality feedback, drive teachers' professional development decisions, facilitate difficult conversations with under-performing teachers, and manage the logistics of conducting classroom observations, as well as pre- and post-observation conferences (Jerald, 2012; Sartin et al., 2011). Rosenholtz (1985) argues that principals play a key role in developing teacher certainty—a necessary pre-condition for teacher effectiveness—by supporting teachers to establish instructional goals and monitoring their progress toward these goals.

Respondents overwhelmingly agreed that an effective teacher evaluation system informs and drives teacher professional development. The educational literature reminds us that VAMs may be valuable in ranking teach-

ers relative to their peers, but they are not designed to provide information about why or how teachers may improve their practice (Curtis & Wiener, 2012). The literature recommends that professional development be tied to evaluation results for all teachers, rather than singling out teachers whose evaluation results are negative (NCTQ, 2011).

The literature indicates that effective professional development requires careful planning and dedicated resources. Joyce and Showers (2002) found that providing opportunities for teachers to practice new skills, as well as coaching during application in authentic classroom contexts, was essential for educational innovations to transfer from the workshop to the classroom. While many models of professional development lack effectiveness as proven by research (Jerald, 2012), the Joyce and Showers model is one of the few evidence-based models (Kuijpers, Houtveen, & Wubbels, 2010).

Finally, survey respondents demonstrated their highest levels of concern regarding tying the results of teacher evaluation to increased compensation for teachers. Increased compensation has been argued as a strategy to attract new teachers and retain highly effective teachers in systems where teacher evaluation data are central to policies aimed at systematically removing the most ineffective teachers (Hanushek, 2009). Proponents also suggest basing tenure decisions (Gordon, Kane, & Staiger, 2006) and staff reduction policies (Boyd, Lankford, Loeb, & Wyckoff, 2010) on teacher evaluation data. Though advocates of teacher deselection acknowledge the potential superiority of systems that increase the effectiveness of individual teachers, they point out that many professional development efforts lack effectiveness and/or scalability. Still others argue for a balance of policies to remove ineffective teachers along with systematic efforts to improve the effectiveness of all teachers (Jerald, 2012). In Indiana, however, educational leaders question whether sufficient resources will be available to implement either strategy.

EMPLOYING A PROCESS FOR IMPLEMENTATION

The development of evaluation models is only part of the challenge. Creating a process with identified elements and protocols is critical to successful implementation. The NCTQ notes that policymakers and education leaders must take steps to “address the anxieties a new evaluation system creates for teachers” (National Council on Teacher Quality, 2011, p. 36). The National Comprehensive Center

for Teacher Quality outlines key components for consideration in developing evaluation plans (Goe, Holdheide, & Miller, 2011) and the Indiana Teacher Appraisal and Support System (IN-TASS) provide examples of a system of training and facilitation to ensure school districts not merely comply with the law, but rather, construct quality systems that ensure equitable, effective, and efficient plans for evaluating educators. The components of this process are noted in Table 8 and are discussed further in the Policy Perspectives on pages 12-14. They include:

1. An understanding of intent and philosophy of teacher appraisals. The purpose of this component of the plan is to ensure a shared belief system on the purpose of teacher evaluations and to ensure transparency on the varied decision points regarding the plan. District teams should develop a strategic communication plan that outlines core belief statements, a process for gathering feedback from key stakeholders, and a clear process for disseminating information.
2. Designing the specific elements of the evaluation system. The specifics of the evaluation system that teachers and principals will be using require collaborative discussion regarding the protocols and procedures of implementation. This discussion should provide district teams an awareness of the various options for observation rubrics, student assessments, weights of measures, observation protocols, and considerations for unique personnel. The discussion should move the group to a consensus on the details for their local plan.
3. Data integrity and transparency. Considerations for either purchasing or adapting an existing data system should be discussed. Part of this discussion should include how the data will be collected, used, stored, and validated. Efficient and effective management of the data from the evaluations will help create trust in the evaluation system and the district.
4. Professional development. Any system for evaluating teachers is only effective if it can tie to professional development needs, both at the individual level and at the district/school level. Ensuring that evaluators have the necessary training to implement the system is important and teachers must also have a sense of certainty around the system. District plans should identify ways that new and struggling teachers will be provided support.

Table 8. Components for a Quality Appraisal Plan

Plan Component	Key Elements
Understanding of Intent and Philosophy of Teacher Appraisals	<ul style="list-style-type: none"> • Purpose of Teacher Evaluations • Ensuring Stakeholder Investment • Strategic Communication Plan <ol style="list-style-type: none"> 1. Core Belief Statements for Teacher Evaluation 2. Process for Gathering Feedback 3. Process for Disseminating Information in Development and Implementation Stages
Elements of the Evaluation System	<ul style="list-style-type: none"> • Legislative Requirements • Scope of System (For whom does the system apply?) • Weight of Measures <ol style="list-style-type: none"> 1. Percentage for Student Learning Data 2. Percentage for Instructional Process Data • Selecting Instructional Process Measure <ol style="list-style-type: none"> 1. Observation Process 2. Standards for Evidence • Selecting Student Learning Measures <ol style="list-style-type: none"> 1. Fundamental Principals of Quality Assessments 2. Process for Selecting and Creating Measures 3. Considerations for Varied Structures and Circumstances • Converting Measures to Teacher Ratings • Scoring/Summative Conference • Definitions and Forms
Data Integrity and Transparency	<ul style="list-style-type: none"> • Process for Collecting, Analyzing, and Storing Data • Linking Data to Teachers • Data Validation • Use of Data
Professional Development	<ul style="list-style-type: none"> • Training for Evaluators and Teachers • Plan to Support New and/or Struggling Teachers • Linking Evaluation Data to District Professional Development
Evaluating the System	<ul style="list-style-type: none"> • Plan for Gathering Feedback • Data Analysis of District/School Evaluation Data • Process to Resolve Discrepancies/Anomalies

Note: Adapted for IN-TASS from Goe, Holdheide, & Miller. (2011). *A practical guide to designing comprehensive teacher evaluation systems*. Washington, DC: National Comprehensive Center for Teacher Quality.

5. Evaluating the system. As school districts evolve through the implementation of their plans, there should be a clear way to evaluate how well the plan is working for teachers and administrators. This would include establishing an on-going review committee to discuss and resolve any anomalies, inconsistencies, and discrepancies in the evaluation system.

As discussed throughout this brief, bringing cultural as well as structural changes to Indiana’s teacher evaluation system is necessary. Engaging key stakeholders in critical reflection and inquiry ensures that those who must implement a new evaluation system own each decision 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. If successful implementation and sustained change are to occur, clarity of purpose and a willingness to accept responsibility for that purpose are

critical (DuFour & Eaker, 1998)

The state of Indiana and those who educate Indiana’s children have an opportunity to “get it right.” Clearly, the state has become a leader in teacher evaluation legislation and policy. Future years will decide if they are a leader in the implementation of a new teacher evaluation system that improves student learning and teacher effectiveness.

CONCLUSIONS AND RECOMMENDATIONS

Conclusion

Employing a Process

The superintendent survey shows that 80% of the districts have adopted a teacher evaluation model. What is not clear (because the survey did not request this information) is whether a process was used to adopt the model. Simply being compliant and adopting a teacher evaluation model to comply with the legislation does not guarantee that the internal norms around teacher quality, student growth and learning, and quality assessment have been created. In other words, simply being compliant will not ensure that the system of evaluation will have created a shared understanding and acceptance among school personnel, a culture of shared responsibility, or a system of internal accountability. Rather, it will only be a system for external accountability.

Recommendation

If a district has not yet decided on a model for their district evaluation plan, it is recommended that they consider employing the process discussed in this brief to determine the best plan for their district. If a district has adopted a model and will be implementing it in the coming school year, it is recommended that they review their plan over the course of the first year of implementation, using the process outlined in this brief. Teacher evaluation plans can be revised at any time. Therefore, engaging district stakeholders in a review of the adopted model and a process to create any missing components of a quality plan should be considered.

Conclusion

Linking Professional Development to Teacher Evaluation

A quality evaluation system should have the capacity to reliably identify effective and ineffective teachers. Additionally, ensuring that teacher ratings can reliably detect teacher strengths and weaknesses is essential for accurately targeting professional development. Evaluation results should be used to identify individual, school, and district-wide needs; target professional learning; gauge teacher growth; and identify potential master teachers who could serve as mentors to new

teachers. Providing job-embedded, ongoing, individualized professional learning is necessary for teacher evaluation to have positive impacts on teacher practice (Goe, Holdheide, & Miller, 2011).

Recommendations

1. Results from all aspects of the teacher evaluation system should be used to inform the district and school leadership on professional development needs that are individual teacher-focused. Districts should no longer engage in random acts of improvement; rather, resources and professional development plans should all be linked to the teacher evaluation data.
2. The state must ensure that resources are available to districts to provide the kinds of quality professional development necessary to ensure that teacher evaluation plans are implemented with fidelity. For example, resources might include financial resources to districts, statewide training, resource materials, and professional interactions on Learning Connections.

Conclusion

Student Learning Assessments

Regardless of the model a district adopts, it has become clear that time, energy, and expertise will be needed to develop the necessary assessments and measures to be used to determine a teacher's student learning rating. Assessment literacy must be developed among teachers and administrators so that confidence and trust in the system can be established and sustained. Teachers and administrators will need to gain greater insight into educational assessments, particularly those that assess student growth. They will need to know and understand how to create assessments and be conversant with the wide array of possible assessment options. And, because the use of multiple measures is recommended in a quality teacher evaluation system, districts will need to have a varied menu of quality assessments for teachers to use. The state of Indiana and local school districts should understand that developing assessment literacy among Indiana's educators will be a developmental process, one that may take several years of work.

Recommendations

1. Districts should have a plan in place, regardless of their timeline for implementing their new teacher evaluation plan, to create the knowledge base and expertise to develop and adopt assessments that can measure student learning.
2. The Indiana Department of Education should work to support local districts with the development of quality assessments for use across grade levels and content areas, as well as provide statewide professional development to help educators become more assessment literate.

Conclusion

Evaluation of Teacher Evaluation Systems

Because the results of teacher evaluations will be used to make personnel and compensation decisions, the fidelity of implementation across the state is critical. As mentioned earlier in this brief, the state of Indiana has given a great deal of flexibility to local districts as they create and develop their teacher evaluation plans. Indiana did not mandate a particular model; rather they provided guidelines for evaluation plans while allowing districts the flexibility in adapting those guidelines locally to develop their own unique plans. On the one hand, this flexibility and local control allows districts to align their plans with local needs and cultures. On the other hand, allowing districts to determine their own unique plans may result in plans that will vary in quality and will be implemented with varied degrees of integrity. One outcome of the legislation could be that districts will submit their plans and simply be compliant with the legislative requirements. Another, more hopeful outcome is that districts will not only be compliant, but will have quality plans in place, implement those plans with fidelity, and monitor the data to continually improve their evaluation systems. In the end, the ultimate goal of teacher evaluation is to improve teaching and learning.

Recommendations

1. It is recommended that a rubric be created that includes the processes and components identified in this brief. This rubric could be used to evaluate district plans that are making a difference, to inform future guidance, and to determine what practices are effective and should be scaled up.

2. Districts should have a clear process in place to evaluate their local teacher appraisal system. A data infrastructure should be in place to monitor student learning and teacher practice ratings district-wide. The data infrastructure should be used to collect, validate, interpret, track, and communicate teacher performance data to inform stakeholders, guide professional development, and assess the system as a whole. Data should be reviewed to determine inconsistencies, and reviewed by an oversight committee. Stakeholder and participant feedback should be solicited and reviewed.

Final Conclusion

The legislation has been passed, a year of guidance and model development has taken place, and soon districts will begin to submit their teacher evaluation plans to the state. Quality implementation is critical. Sustained support from the state is imperative. Flexibility to modify and adjust both state policy and district plans must be ensured. According to the MET Project, 2012, new approaches to measuring teacher effectiveness, while not perfect, significantly outperform traditional measures. The conversation on teacher evaluations in Indiana has shifted; the next few years will determine if the conversations have led to improved instruction and student learning.

Policy Perspective

The Evanston Teacher Appraisal System: Outcomes and Lessons Learned

Dr. Hardy Murphy

During the 2007-08 negotiations, my district needed to re-submit our teacher appraisal system in order to be in compliance with the Illinois state school code. As a result, the administration and teachers' union engaged in a collaborative process to redesign our appraisal process, with the understanding that the new system would include student growth and achievement as 50% of the teacher's rating. This system was first implemented in the 2008-09 school year.

Persistent themes defined the dialogue accompanying the appraisal system's development. The early discussions centered on a few predictable concerns. Is it fair? Is it helpful? Can it be done? Why us? Additionally, whether true academic achievement and growth could be accurately measured and whether or not the student's overall life experiences could be overcome in the teaching and learning process were intensely debated. There was conversation concerning the feasibility of a system with differential expectations, applications, and implementation standards for professionals with different roles in the teaching and learning process. The debate over the relative validity and reliability of teacher-made measures vs. standardized tests, performance measures and rubrics, and teacher portfolios generated thought-provoking commentary. All of these concerns had to be addressed in the context of a belief that teaching is about something more than testing; it is about supporting the academic development of students in their journey to becoming well-adjusted adults who participate constructively in their community.

A breakthrough occurred when a series of very simple beliefs and questions were posed: 1) Do we believe that being in a teacher's classroom for a year of instruction is of benefit to a student?

2) Shouldn't we be able to demonstrate the difference a year's worth of teaching makes for a student? 3) Shouldn't parents expect a year of school to make a difference? 4) If we show more than a year's worth of improvement for a student, shouldn't the teacher be recognized for it? 5) If a student makes less than a year's worth of growth, don't we feel this to be unsatisfactory? 6) If less than a year's worth of growth is made after a year of instruction shouldn't we provide an assessment and a plan of action? 7) Shouldn't all teachers involved in the instructional process for a child share in the responsibility and the accountability for the child's progress? Every teacher involved understood and believed that they were responsible for and could answer for a year's growth. Once this concept was understood, then creating the other components of the system was possible.

As a result of a system that links instructional processes to student learning outcomes, an understanding of the relationship between the strengths and weaknesses of the district's instructional program is occurring at all levels. Teachers are reviewing the strengths and areas for improvement within classrooms. Principals are identifying areas that are strongly associated with student performance in their schools to focus staff efforts. The identification of highly correlated trends between instructional features and student outcomes is informing district professional development and data management. From a governance perspective, it has highlighted a need to align board expectations and goals with achievement and growth expectations in the teacher appraisal system.

Over the first two years of implementation, the appraisal system has yielded encouraging results. In the year prior to the creation of the system, virtually all of the district's teachers were awarded the highest ratings (96%). Virtually none of our teachers were rated as "needs to improve" (2%), or unsatisfactory (2%). In the first year of implementation, this changed to 70% excellent, 28% needs to improve, and 2% unsatisfactory. At the conclusion of the second year, there was an increase in the number of teachers achieving the highest rating of excellent to 76%. However, I feel that this is good and a natural outgrowth of an appraisal model that is focused upon a constructive problem-solving conversation between teachers and principals about student achievement and how to improve teaching. Improved student performance on the measures used to document growth and achievement over the initial years of implementation only validates my belief and confirms that what we have accomplished is the creation of an appraisal system that is helpful for both students and teachers.

The resolution of initial concerns in the development phase and what has been learned over a 3-year period in the implementation of our appraisal system offer insights into the belief systems of teachers and administrators and how a profession whose judgments of practitioner effectiveness historically anchored in the security of instructional inputs and teacher processes may successfully venture into the challenging world of student learning outcomes and their measurement.



Dr. Hardy Murphy

Evaluation Results from Two Years of Implementation

2010	Unsatisfactory	Satisfactory	Excellent
Achievement	2.07%	14.50%	83.43%
Danielson	2.66%	40.83%	56.51%
Final Summative	2.07%	27.81%	70.12% (237)
2011	Unsatisfactory	Satisfactory	Excellent
Achievement	2.18%	8.25%	89.56%
Danielson	1.46%	33.01%	65.53%
Final Summative	1.21% (5)	23.06% (95)	75.73% (312)

Dr. Hardy Murphy, Superintendent, Evanston-Skokie School District 65, Evanston, Illinois. IN-TASS Consultant.

Policy Perspective

Redesigning Teacher Evaluation Systems: Lessons for States

Lynn Holdheide

The accelerated interest in the redesign of teacher evaluation systems can be attributed, at least partially, to the current era of accountability for a high standard of student, educator, and school performance. This, along with the growing body of evidence that cites teachers as the most influential school-based factor on student achievement, coupled with the various initiatives (ESEA Flexibility Waivers) and funding streams (e.g., Race to the Top, School Improvement Grants, Teacher Incentive Funds), have incentivized states and districts to focus on the identification and retention of highly effective teachers. Accordingly, reform agendas have increasingly come to focus on assessing, developing, and supporting teacher effectiveness through the strengthening of teacher evaluation systems.

Redesigning teacher evaluation has the potential to transform the profession by ensuring that the highest quality of instruction is occurring in every classroom. But reforms will not have this kind of impact unless they are designed to continually improve teacher capacity so that all students are provided an opportunity to be college and career ready. The TQ Center has partnered with state education agencies to retool teacher evaluation so that it becomes a meaningful experience, not just a pointless exercise. States would do well if they understand the lessons that are emerging:

- **Take the time to reach a consensus on the goal and purpose of the system.** All too often, states and districts are quick to adopt a particular measure or rubric without taking the time to determine the evaluation system goals and purpose. Goal development helps to guide and focus design decisions and, when integrated within other state or district initiatives, can provide a greater sense of coherence and credibility among stakeholders.



Lynn Holdheide

- **Secure stakeholder buy-in and support.** Promoting educator voice is imperative so that the evaluation model is perceived as fair, accurate, and useful for the stakeholders it is designed to support. Building a system whereby educators consider the process equally advantageous, and not as happening to them, will go a long way in gaining stakeholder support and improving teacher capacity and student outcomes.
- **Ensure coherence and alignment across the teacher career continuum.** Although teacher evaluation plays a vital role in improving teacher capacity and student performance, a focus on efforts to improve educator effectiveness at every stage of the career pipeline also matters a great deal. Effective performance management requires a data-driven system that informs human capital management decisions all built within teacher evaluation frameworks and coupled with increased opportunities for professional learning and formative feedback. Creating coherence between educator recruitment, preparation, licensure, induction, mentoring, professional development, compensation, and human resource management is at the heart of recruiting, developing, and retaining high-quality educators.
- **Focus on Instruction.** Implementation of the College and Career Readiness Standards (CCRS) requires that standards are translated into an aligned, universally designed curriculum; teachers deliver instruction that is differentiated and designed to elicit high cognitive demand and levels of student engagement; and assessments are aligned and accessible so that knowledge of the standards can be fully demonstrated. This era of reform requires a change in instructional practices and a greater understanding of what effective teaching and learning looks like in the 21st Century. Therefore, the credibility (and validity) of teacher evaluation rests on the system's ability to provide teachers with the competencies to employ these instructional shifts with fidelity, to accurately identify the individual development needs of teachers, and to promote independent and shared reflection that enable teachers, through the provision of sustained

professional learning and support, to continually evaluate and improve their practice.

- **Continually Gauge Progress and Validate Efforts.** It is increasingly clear that states are compelled to retool teacher evaluation at a rate in which guidance from practice and research is not yet available. The way forward, therefore, is to build a system of monitoring statewide implementation that measures progress toward the identified goals. In particular, the pilot projects would be structured to collect and use data to determine the efficacy of state and district efforts and initiatives (e.g., teacher preparation, mentoring, job-embedded professional development) through analysis of equitable distribution, improved teacher capacity, and student growth. Likewise, given the complexity of incorporating measures of student growth, the need to commission studies that coincide with implementation can help determine technical adequacy (e.g., reliability and validity) in making high-stakes decisions.

¹ Rivkin, S.G., Hanushek, E.A., & Kain, J.G. (2005). Variable definitions, data, and programs for 'teachers, students, and academic achievement.' *Econometrica Supplementary Material*, 73(2); Leithwood, K., Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning* (Review of research). New York: The Wallace Foundation. Retrieved September 24, 2011, from <http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Documents/How-Leadership-Influences-Student-Learning.pdf>

² Coleman, D., & Pimentel, S. (2012). *Publishers' criteria for the Common Core State Standards in English language arts and literacy, grades 3-12*. Retrieved online July 7, 2012 at: <http://www.achievethecore.org/downloads/Publishers%20Criteria%20for%20Literacy%20for%20Grades%203-12.pdf?20120412>; Conley, D.T. (2011). Building on the common core. *Educational Leadership*, 68(6), 16-20; Council of Chief State School Officers. (2011, April). *Interstate Teacher Assessment and Support Consortium (InTASC) Model Core Teaching Standards: A Resource for State Dialogue*. Washington, DC: Author.

Lynn Holdheide is a Research Associate at Vanderbilt University / National Comprehensive Center for Teacher Quality, www.tqsource.org

Policy Perspective

Training Facilitates Development of Local Evaluation Framework

Ryan Snoddy

Three issues stood out when teacher evaluation discussions began in the Northwestern School Corporation. While legislative action of the state required all Indiana school corporations to develop new evaluation models, our local committee viewed this challenge as an opportunity to improve teacher effectiveness and student academic growth. First we accepted the new requirements of the Indiana legislation. Secondly, the past evaluation process of the school corporation was developed through research and was previously rubric driven, and that framework served as the foundation for the initial discussions in the development of our new model. Finally, it was important to develop the new framework with sincere participation and input from all teachers and administrators. We acknowledged that we couldn't simply adopt a "model", but rather, we needed to engage key stakeholders in conversations to ensure the creation of a culture of shared responsibility throughout the development process of the new evaluation model.

Step one in our local process began in September 2011 when a local Northwestern Teacher Evaluation Committee (NorTEC) consisting of eight teachers, six principals, and two central office administrators was established. Three of those committee members had the opportunity to hear Dr. Hardy Murphy discuss the Evanston/Skokie Professional Appraisal System at Indiana University Kokomo. The process and the model presented resonated with committee members. When the opportunity for Northwestern to participate in a similar process arose, the Northwestern committee was on board.



Ryan Snoddy

The task of developing a local teacher evaluation framework can be overwhelming. However, with a clear, organized process, the work can be manageable. Using the Evanston framework and support from the Indiana Teacher Appraisal and Support System (IN-TASS) we began to take the necessary steps to create and develop a quality plan. This process allowed our committee to engage in critical conversations about our beliefs on teacher evaluation, our process for gathering and sharing important feedback and information, the pros and cons of various evaluation models and the importance of collegial decision-making. The primary focus of our first NorTEC meeting was to develop a sound purpose that serves as a foundation and supports our new evaluation system. That purpose, "... enhance quality instruction, encourage professional growth of the teacher, and improve student achievement" is revisited at the beginning of each committee meeting and has often been the guiding principle of on-going discussions in the development of our evaluation system. The development of our purpose has also challenged our committee to consider the results of classroom instruction, improved student achievement, and the compelling indicators of teacher effectiveness.

Employing a process has been instrumental in helping guide the NorTEC group in the development of the components of our teacher appraisal plan. The following statement from our high school principal, Al Remaly, summarizes the benefit:

[Using a process] has allowed us to define our core beliefs for student and teacher expectations. The collaboration we have had has brought forth a new perspective on teacher evaluation. It has allowed us to have very open and honest discussions regarding our thoughts

on the qualities that a teacher needs to be exemplary. I am confident that, at the completion of this process, we will have an evaluation model that is very effective because of the collaboration that went into the development of our local plan.

As Northwestern School Corporation proceeds towards implementation of our teacher appraisal system we still have much to undertake. The final document has been reviewed and approved by all teachers and presented and adopted by the Board of School Trustees and will be submitted to the Indiana Department of Education this summer. However, the master agreement between Northwestern School Corporation and the teachers' association contains evaluation language and therefore the parties are bound by contract language to comply with the current evaluation model. Discussion has been held to consider the possibility of piloting the Danielson teacher effectiveness rubric for the 2012-13 school year, while details are finalized with regard to the student growth component of the appraisal system. We recognize that this is just the beginning of our journey; our conversations will continue and we will review data from our plan to ensure fidelity of implementation.

Dan Robinson, President of the Northwestern teachers' association summarizes our work:

[A defined process] has provided sound fundamental guidance and suggestions in a time of uncertainty. We were able to sort through the vast variety of options open to schools and select those with the most merit for inclusion into a new plan.

Ryan Snoddy, Superintendent, Northwestern School Corporation, Kokomo, Indiana, and Member IN-TASS Leadership Cadre.

AUTHORS

Cassandra M. Cole

(cmcole@indiana.edu) is Director of the Center on Education and Lifelong Learning at Indiana University, and leads the IN-TASS project.

Jim Ansaldo

(jansaldo@indiana.edu) is a Research Associate at the Center on Education and Lifelong Learning at Indiana University. Jim recently completed his Ph.D. in Curriculum Studies at Indiana University and served on the state Leadership Cadre for IN-TASS.

James N. Robinson

(jarobins@indiana.edu) is an Ed.D. candidate in Education Leadership and Policy at Indiana University. He is a Research Associate at the Center on Education and Lifelong Learning, Indiana University and served on the state Leadership Cadre for IN-TASS.

Rod S. Whiteman

(whiteman@indiana.edu) is a Ph.D. student in Education Policy Studies at Indiana University. His research focuses on schools as organizations, and education policy impacts on school leaders, teachers, and the teaching profession.

Terry E. Spradlin

(tspradli@indiana.edu) is the Director for Education Policy and HSSSE at the Center for Evaluation & Education Policy.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. Hardy Murphy, Superintendent, Evanston-Skokie School District 65, Evanston, Illinois, for his willingness to provide his support to various Indiana school districts, his leadership in helping to create the process described in this brief, and for his submission of his Policy Perspective letter. Additionally, the authors are grateful to Ryan Snoddy, Superintendent for Northwestern School District, and Dr. Lynn Holdheide, Research Associate at Vanderbilt University / National Comprehensive Center for Teacher Quality, for their submissions of Policy Perspective letters.

The authors would also like to thank Dr. John Ellis, former Executive Director of the Indiana Association of Public School Superintendents, for his support with the superintendents' survey, and we thank the many superintendents who completed the survey.

REFERENCES

- Bill and Melinda Gates Foundation. (2012). *Gathering feedback for teaching: Combining high quality observations with student surveys and achievement gains*. Seattle, WA: Bill and Melinda Gates Foundation, The MET Project.
- Boyd, D.J., Lankford, H., Loeb, S., & Wyckoff, J.H. (2010). *Teacher layoffs: An empirical illustration of seniority vs. measures of effectiveness*. Washington, DC: The Urban Institute.
- Coggs, J.G. (2007). *Communication framework for measuring teacher quality and effectiveness: Bringing coherence to the conversation*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Curtis, R., & Wiener, R. (2012). *Means to an end: A guide to developing teacher evaluation systems that support growth and development*. Washington, DC: The Aspen Institute.
- Daley, G., & Kim, L. (2010). *A teacher evaluation system that works*. National Institute for Excellence in Teaching Working Paper. Santa Monica, CA: National Institute for Excellence in Teaching.
- Danielson, C. (2007). *Enhancing professional practice: A framework for teaching* (2nd ed.). Alexandria, VA: ASCD.
- Darling-Hammond, L., Wise, A.E., & Pease, S.R. (1983). Teacher evaluation in the organizational context: A review of the literature. *Review of Educational Research*, 53(3), 285-328.
- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practice for enhancing student achievement*. Bloomington, IN: National Education Service.
- Elmore, R.F. (2004). *School reform from the inside out: Policy, practice, and performance*. Cambridge, MA: Harvard Education Press.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Goe, L., Holdheide, L., & Miller, T. (2011). *A practical guide to designing comprehensive teacher evaluation systems*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Gordon, R., Kane, T.J., & Staiger, D.O. (2006). *Identifying effective teachers using performance on the job*. Washington, DC: The Brookings Institution.
- Hanushek, E.A. (2009). Teacher deselection. In D. Goldhaber & J. Hannaway (Eds.), *Creating a new teaching profession* (pp. 165-80). Washington, DC: Urban Institute Press.
- Indiana Department of Education (IDOE). (2012). *Legislative guidance: Rigorous measures*. Indianapolis: Author.
- Indiana Department of Education (IDOE). (2011). *RISE evaluation and development system: Evaluator and teacher handbook*. Indianapolis: Author
- Indiana Senate Enrolled Act 001 (Public Law 90), Jerald, C.D. (2012). *Movin' it and improvin' it! Using both education strategies to increase teaching effectiveness*. Washington, DC: Center for American Progress.
- Joyce, B.R., & Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Kuijpers, J. M., Houtveen, A.A.M., & Wubbels, T. (2010). An integrated professional development model for effective teaching. *Teaching and Teacher Education*, 26(8), 1687-94.
- Malen, B. (1995). The micropolitics of education: Mapping the multiple dimensions of power relations in school policies. In J.D. Scribner & D.H. Layton (Eds.), *The study of educational politics* (pp. 147-67). London: The Falmer Press.
- Marzano Research Laboratory. (2011). *The Marzano teacher evaluation model*. Englewood, CO: Marzano Research Laboratory (marzanoresearch.com).
- Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education*, 79(4), 33-53.
- Morris, S. (2012, March 23). Use the data, but constructively. *The New York Times*. Retrieved from <http://www.nytimes.com/roomfordebate/2012/01/16/can-a-few-years-data-reveal-bad-teachers/use-the-data-but-use-it-constructively>
- National Council on Teacher Quality. (2011). *State of the states: Trends and early lessons on teacher evaluation and effectiveness policies*. Washington, DC: Author.
- The New Teacher Project. (2010). *Teacher evaluation 2.0*. New York: Author. Retrieved March 21, 2012, from <http://www.tntp.org/files/Teacher-Evaluation-Oct10F.pdf>
- The New Teacher Project. (2011). *Rating a teacher evaluation tool: Five ways to ensure classroom observations are focused and rigorous*. New York: Author. Retrieved March 21, 2012, from http://tntp.org/assets/documents/TNTP_RatingATeacherObservationTool_Feb2011.pdf
- Newmann, F., King, B., & Youngs, P. (2000). *Professional development that addresses school capacity*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Rosenholtz, S.J. (1985). Effective schools: Interpreting the evidence. *American Journal of Education*, 93(3), 352-388.
- Rosenholtz, S.J. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Sartain, L., Stoelinga, S.R., & Brown, E.R. (2011). *Rethinking teacher evaluation in Chicago: Lessons learned from classroom observations, principal-teacher conferences, and district implementation*. Chicago: Consortium on Chicago School Research.
- Whiteman, R.S., Shi, D., & Plucker, J.A. (2011). *Revamping the teacher evaluation process*. Bloomington, IN: Center for Evaluation and Education Policy.

WEB RESOURCES

National Council on Teacher Quality

<http://www.nctq.org/p/>

The New Teacher Project

<http://tntp.org/>

National Institute for Excellence in Teaching

<http://niet.org/>

National Comprehensive Center for Teacher Quality

<http://www.tqsource.org/>

The Danielson Group

<http://www.danielsongroup.org/>

National Center for the Improvement of Educational Assessment

<http://www.nciea.org/>

Indiana Department of Education RISE Evaluation and Development System

<http://www.riseindiana.org>

The MET Project

<http://www.metproject.org/>

Education Policy Briefs are executive edited by Jonathan A. Plucker, Ph.D. and published by the

Center for Evaluation & Education Policy

Indiana University

1900 East Tenth Street

Bloomington, IN 47406-7512

812-855-4438

More about the Center for Evaluation & Education Policy
and our publications can be found at our Web site:

<http://ceep.indiana.edu>



0829120940