

Stated Briefly

The utility of teacher and student surveys in principal evaluations: An empirical investigation



Stated Briefly

Keke Liu
Jeff Springer
David Stuit

Basis Policy Research

Jim Lindsay
Yinmei Wan

American Institutes for Research

In collaboration with the Midwest Urban Research Alliance

This brief summarizes the findings of a recent study that examined whether adding results from teacher and student feedback surveys to one district's principal evaluation system strengthened the link between principals' evaluation ratings and their schools' achievement gains in math and reading. The study also examined whether certain subscales from student and teacher surveys were better than others at improving the link between evaluation system ratings and schoolwide student achievement. The study found that adding the surveys to the district's principal evaluations strengthened the link between the evaluation ratings and schoolwide student achievement in math and in a math-reading composite but not in reading alone.

This brief summarizes the findings of Liu, K., Springer, J., Stuit, D., Lindsay, J., & Wan, Y. (2014). *The utility of teacher and student surveys in principal evaluations: An empirical investigation* (REL 2015–047). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. That report is available at <http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=357>.



Institute of Education Sciences
U.S. Department of Education



Regional Educational Laboratory
At American Institutes for Research

Why this study?

State and local education agencies across the Regional Educational Laboratory Midwest Region (which includes Illinois, Indiana, Iowa, Michigan, Minnesota, Ohio, and Wisconsin) are looking for ways to improve their principal evaluation systems to help provide better feedback to school principals on how their leadership practices influence student achievement. Previous research shows that principals have significant effects on student achievement but that the effects come indirectly through their role in cultivating school conditions that promote student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004). Evaluating these indirect paths of influence therefore requires sound measures of the school conditions that effective principals create to improve student achievement.

In partnership with the Midwest Urban Research Alliance, the study team examined whether adding student and teacher feedback surveys to an existing evaluation system strengthened the statistical link between principals' evaluation ratings and their schools' average student achievement gains in math and reading. If surveys assessing qualities of the school that are under principals' control can strengthen the statistical link between principals' evaluation ratings and their schools' achievement gains, then incorporating the survey results into principal evaluation systems could improve the quality of feedback to principals by providing evidence on how well they are cultivating school conditions that facilitate learning. The study also examined whether certain domains of the students and teacher surveys (box 1) were better than others at improving the link between evaluation system ratings and schoolwide student achievement.

The study was conducted using data from the 2011/12 school year on principals in 39 elementary and secondary schools located in a medium-size urban district in the Regional Educational Laboratory Midwest Region.

Box 1. Data and methods

The data included:

Existing principal evaluation measures. The district uses a principal evaluation system that has two measures: supervisor ratings of principals and schoolwide average student attendance rates.

- **Supervisor ratings.** The district uses two rubrics that supervisors complete following their observations of the principal. The rubrics represent core competencies in two areas: job function competency and leadership skill competency.
- **Attendance.** This measure represents the schoolwide average of student attendance rates.

Tripod student and teacher surveys. These surveys were developed by Harvard University researcher Ronald Ferguson for districts to administer to students, teachers, and principals in order to obtain information about the learning environment within classrooms, schools, and districts.¹ The study team examined Tripod survey data collected by the district from 8,345 students in grades 3–11 and from 541 teachers. The surveys consist of statements—or items—to which the student or teacher responds with “totally untrue,” “mostly untrue,” “somewhat true,” “mostly true,” or “totally true.” The items in each survey are grouped into domains that measure distinct conditions of the schools' teaching and learning environment.

- **The Tripod Student Perception Survey.** This survey measures students' perceptions of their classroom and school environments. It takes approximately 30 minutes to complete. Two measures were constructed from the items in this survey and used in the analysis:
 - **Classroom instructional environment.** This measure is based on the average rating of 36 items that are designed to assess seven domains of classroom instruction, including the degree to which students

(continued)

Box 1. Data and methods *(continued)*

perceive that their teacher cares about students, controls behavior, clarifies lessons, challenges students, captivates students, confers with students, and helps students consolidate knowledge.

- **School safety and climate.** This measure is based on the average rating of six survey items that assess the degree to which students feel that they are safe in the classroom and school and that student safety is emphasized by administrators and teachers.
- **The Tripod Teacher Survey.** This survey is designed to capture teachers' perceptions of school conditions that research has shown to influence the quality of teaching and learning. It is designed to take about 30 minutes to complete. Four measures were constructed from the items in this survey and included in the analysis:
 - **Instructional leadership.** This measure is based on teachers' average ratings of 18 items that assess the expertise of school instructional leaders in promoting a climate of learning, managing instruction, and defining the school's mission.
 - **Professional learning community.** This measure is based on teachers' average ratings of 12 items that assess the amount of time spent in professional learning community activities and in collaboration with teachers on curriculum design and assessment.
 - **Quality of professional development.** This measure is based on teachers' average ratings of 14 items that assess the effectiveness of professional development activities and of school leaders' support for professional development.
 - **Cultural press for excellence.** This measure is based on teachers' average rating of three items that assess whether the school has created a culture of holding adults accountable for excellence and of setting and achieving goals.

School achievement growth scores. The study team used value-added analysis to calculate schools' average achievement gains in math, reading, and a math-reading composite based on student-level data for students in grades 3–11 from the 2011/12 school year. The data used in the value-added analysis included test scores from the fall 2011 and spring 2012 administrations of the Northwest Evaluation Association's (2011) Measures of Academic Progress and student background characteristics, including race/ethnicity, gender, English learner status, special education status, and mobility.²

The study team also used an accepted statistical approach (regression analysis) to test whether the surveys and subscales improved the statistical link between principals' evaluation ratings and schoolwide student achievement. Statistically, the link between the evaluation system ratings and schoolwide student achievement is represented by the amount of variation in achievement that is explained by the evaluation system ratings. The more between-school variation that is explained by the ratings (indicated by a percentage), the greater the link between the ratings and achievement.

Notes

1. For more information on the Tripod surveys, see <http://tripoded.com>.

2. Mobility data indicate students who are new to the school in 2011/12 for reasons other than normal grade promotion.

What the study found

The study found that adding the average ratings from student and teacher survey measures to the district's existing principal evaluation system would help improve the link between the evaluation system ratings and school-level achievement in math and in the math-reading composite but not in reading.

Adding average ratings from the entire set of student and teacher survey measures to the existing principal evaluation system significantly improved the link between the evaluation ratings and schools' achievement growth in math and in the math-reading composite but not in reading

Adding average ratings from all six teacher and student survey measures to the district's existing set of principal evaluation measures (supervisor ratings and school attendance rates) significantly increased the evaluation system's link to value-added achievement in math by 39.1 percentage points and to value-added achievement in the math-reading composite by 36.2 percentage points (figure 1). But including average survey ratings did not improve the existing system's link with schools' value-added achievement in reading.

When added individually, three survey subscales produced a statistically significant improvement in the link between the existing principal evaluation system and school-level student achievement in math and in the math-reading composite

Two of the four teacher survey measures (instructional leadership and cultural press for excellence) and one of the two student survey measures (classroom instructional environment) showed a statistically significant positive relationship with school-level achievement growth and collectively improved the link between the principal evaluation ratings and schools' achievement growth in math and the math-reading composite when added to the existing set of evaluation measures. The instructional leadership subscale increased the link between the principal evaluation system and achievement in the math-reading composite by 10.7 percentage points, the cultural press for excellence measure increased the link by 8.1 percentage points, and the classroom instructional measure increased the link by 13.5 percentage points (figure 1).

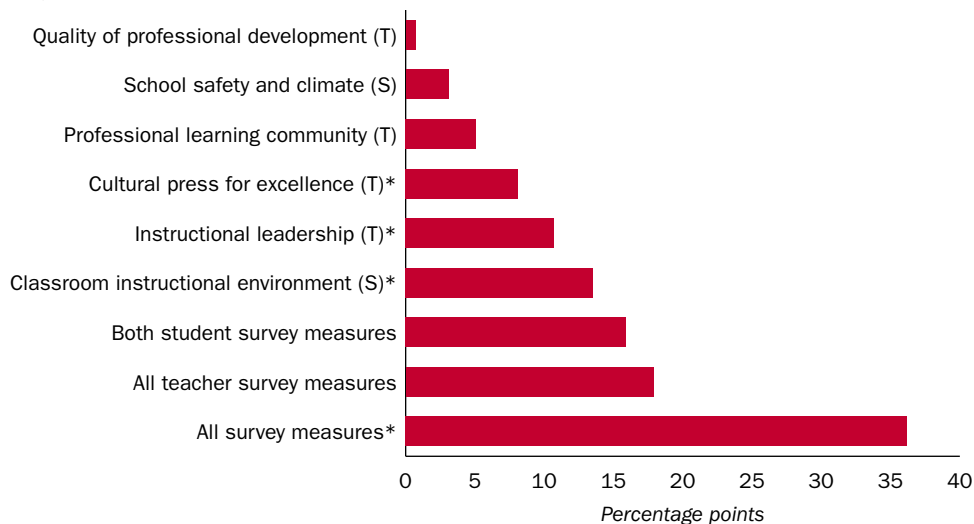
However, when schools' achievement growth in math and reading was considered separately, different patterns of relationships became apparent. The percentage point increase in links between principal evaluation scores and achievement growth in math was 12.1 percentage points when instructional leadership was added to the evaluation system, 8.3 percentage points when cultural press for excellence was added, and 14.5 percentage points when classroom instructional environment was added (all three increases were statistically significant at $p < .10$). Adding the same survey measures to principal evaluation scores yielded no improvements in the link between those scores and schools' achievement growth in reading.

The optimal set of survey measures included the instructional leadership measure from the teacher survey and the classroom instructional environment measure from the student survey

The instructional leadership measure (from the teacher survey) and the classroom instructional environment measure (from the student survey) represented the optimal combination of survey measures. Adding the two measures improved the link between principals' ratings and student achievement gains by 28.8 percentage points in math (14.5 percentage points for classroom instructional environment plus 14.3 percentage points for instructional leadership; figure 2) and by 26.5 percentage points in the math-reading composite (13.5 percentage points for classroom instructional environment plus 13.0 percentage points for instructional leadership; figure 3). Including any of the other subscales on top of these added little to the link between the evaluation system and schoolwide student achievement in math and the math-reading composite.

Figure 1. Two of the four teacher survey measures and one of the two student survey measures significantly improved the link between principals’ evaluation ratings and their school’s composite reading and math achievement gains

Improvement in the link between principal evaluation ratings and school achievement gains in the math-reading composite above and beyond the existing measures



* indicates a statistically significant relationship at $p < .10$.

(T) indicates that the subscale is in the teacher survey; (S) indicates that the subscale is in the student survey.

Note: Based on surveys from 541 teachers and 8,345 students in 39 schools. The bars indicate the percentage of variation in schools’ value-added achievement (math and reading composite) that is explained when each survey measure is added to the existing principal evaluation measures. The more variation that is explained by the ratings, the greater the alignment between the principal evaluation system and schoolwide student achievement.

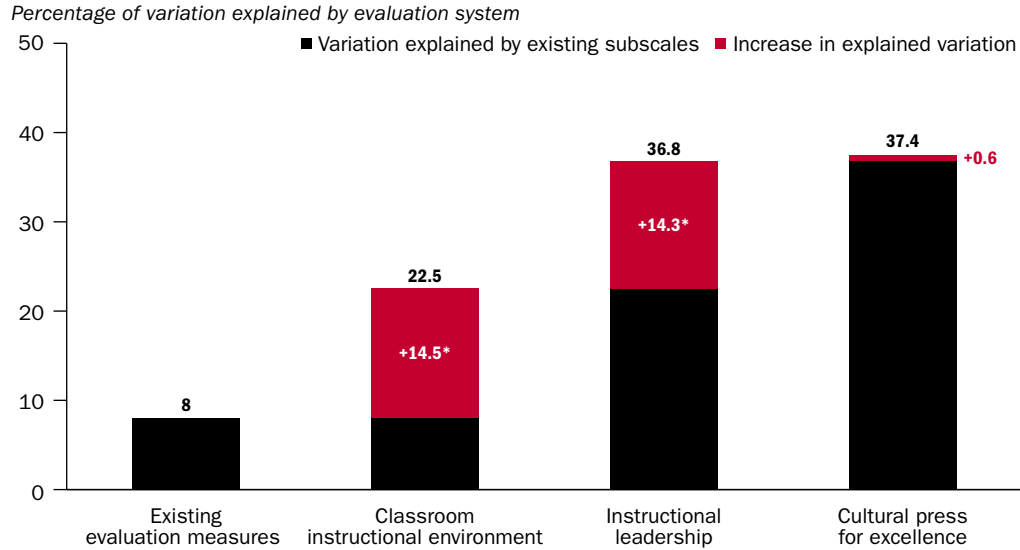
Source: Authors’ analysis based on data provided by the district.

Implications for local and state education agencies

This study found evidence to support the use of student and teacher survey measures in this district’s principal evaluation system. Adding all six survey measures increased the link between principals’ evaluation ratings and their schools’ student achievement gains (math-reading composite) by 36.2 percentage points. These results suggest that surveys are a viable option for districts to consider when developing their principal evaluation systems. However, the benefits of the surveys will depend on the quality of the existing principal evaluation measures.

The findings from this study underscore the importance of incorporating measures of principals’ instructional leadership practices into evaluation systems. Both of the survey measures that were most strongly associated with student achievement gains—instructional leadership from the teacher survey and classroom instructional environment from the student survey—provide a view of the principal as the instructional leader within the school. This finding is consistent with the research literature, which emphasizes that principals play a critical role in shaping the quality of instruction in their schools and, in doing so, can significantly influence student achievement (Béteille, Kalogrides, & Loeb, 2009; Grissom & Loeb, 2009; Portin et al., 2009). Districts that lack sound measures of instructional leadership in their evaluation systems might consider using student and teacher feedback surveys to improve feedback given to principals regarding how they shape the instructional conditions in their schools.

Figure 2. Adding the classroom instructional environment subscale from the student survey and the instructional leadership subscale from the teacher survey improved the link between the principal evaluation system and student achievement in math by 28.8 percentage points

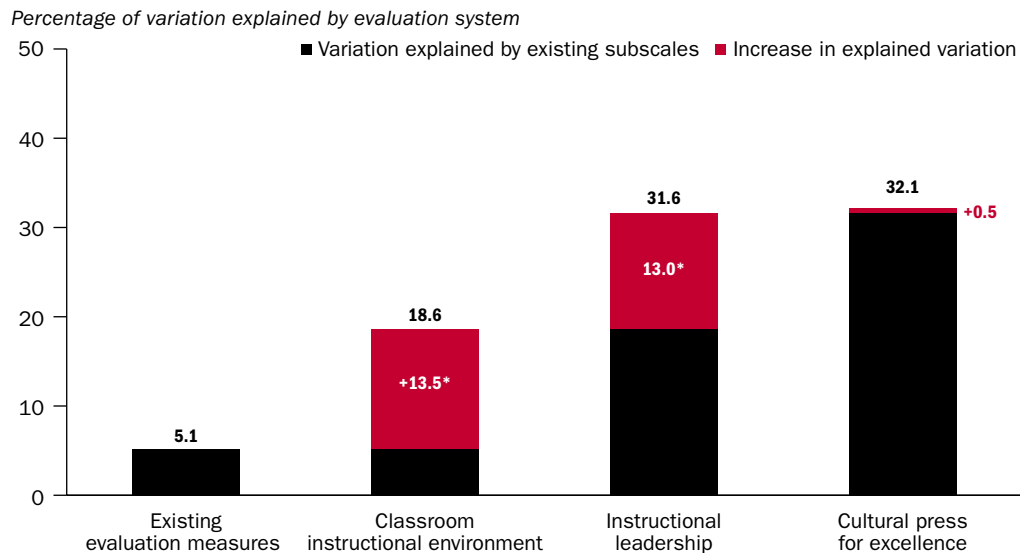


* Increase in explained variance is significant ($p < .10$).

Note: Based on surveys collected from 541 teachers and 8,345 students in 39 schools. Classroom instructional environment is a subscale from Tripod’s student survey; instructional leadership is a subscale from Tripod’s teacher survey. The amount of explained variation indicates the strength of the statistical relationship. The more variation that is explained by the ratings, the greater the alignment between the principal evaluation system and achievement.

Source: Authors’ analysis based on data provided by the district.

Figure 3. Adding the classroom instructional environment subscale from the student survey and the instructional leadership subscale from the teacher survey improved the link between the principal evaluation system and schoolwide student achievement (math-reading composite) by 26.5 percentage points



* Increase in explained variance is significant ($p < .10$).

Note: Based on surveys collected from 541 teachers and 8,345 students in 39 schools. Classroom instructional environment is a subscale from Tripod’s student survey; instructional leadership is a subscale from Tripod’s teacher survey. The amount of explained variation indicates the strength of the statistical relationship. The more variation that is explained by the ratings, the greater the alignment between the principal evaluation system and achievement.

Source: Authors’ analysis based on data provided by the district.

Limitations of the study

The study has two limitations that need to be considered when reflecting on its findings.

First, the findings were based on data drawn from a single urban school district located in the REL Midwest Region. The school district may serve a different student population, may have different principal evaluation policies, or may exhibit a different type of professional culture when compared with other school districts. Therefore, the results of this study do not guarantee that student and teacher surveys will add significant informational value to all districts' principal evaluation systems.

Second, this study examined a limited set of principal performance measures. School districts with validated supervisor rating tools that are strongly correlated to student achievement may find that adding new surveys does not have as much incremental value. The analytic strategy outlined in the full report provides a roadmap for how districts can evaluate the incremental value of new surveys or other new measures that are candidates for inclusion in evaluation systems.

References

- Béteille, T., Kalogrides, D., & Loeb, S. (2009). *Effective schools: Managing the recruitment, development, and retention of high-quality teachers*. Washington, DC: The Urban Institute, National Center for Analysis of Longitudinal Data in Education Research. <http://eric.ed.gov/?id=ED509688>
- Grissom, J., & Loeb, S. (2009). *Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills* (Working Paper No. 35). Washington, DC: The Urban Institute, National Center for Analysis of Longitudinal Data in Education Research. Retrieved June 3, 2015, from <http://www.urban.org/sites/default/files/alfresco/publication-pdfs/1001443-Triangulating-Principal-Effectiveness-How-Perspectives-of-Parents-Teachers-and-Assistant-Principals-Identify-the-Central-Importance-of-Manage-rial-Skills.pdf>.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. Minneapolis, MN: University of Minnesota, Center for Applied Research and Educational Improvement; Toronto, Canada: University of Toronto, Ontario Institute for Studies in Education. Retrieved June 3, 2015, from <http://www.wallacefoundation.org/knowledge-center/school-leadership/key-research/Documents/How-Leadership-Influences-Student-Learning.pdf>.
- Portin, B. S., Knapp, M. S., Dareff, S., Feldman, S., Russell, F. A., Samuelson, C., & Ling Yeh, T. L. (2009). *Leadership for learning improvement in urban schools*. Seattle, WA: University of Washington, Center for the Study of Teaching and Policy. Retrieved June 3, 2015, from <http://www.wallacefoundation.org/knowledge-center/school-leadership/district-policy-and-practice/Documents/Leadership-for-Learning-Improvement-in-Urban-Schools.pdf>.
- Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44(5), 635–674. <http://eric.ed.gov/?id=EJ818931>

The National Center for Education Evaluation and Regional Assistance conducts unbiased large-scale evaluations of education programs and practices supported by federal funds, provides research-based technical assistance to educators and policymakers, and supports the synthesis and the widespread dissemination of the results of research and evaluation throughout the United States.

April 2016

This report was prepared for the Institute of Education Sciences (IES) under Contract ED-IES-12-C-0004 by Regional Educational Laboratory Midwest administered by American Institutes for Research. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This REL report is in the public domain. While permission to reprint this publication is not necessary, it should be cited as:

Liu, K., Springer, J., Stuit, D., Lindsay, J., & Wan, Y. (2016). *Stated Briefly: The utility of teacher and student surveys in principal evaluations* (REL 2016–132). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. Retrieved from <http://ies.ed.gov/ncee/edlabs>.

This report is available on the Regional Educational Laboratory website at <http://ies.ed.gov/ncee/edlabs>.

The Regional Educational Laboratory Program produces 7 types of reports

	Making Connections Studies of correlational relationships
	Making an Impact Studies of cause and effect
	What's Happening Descriptions of policies, programs, implementation status, or data trends
	What's Known Summaries of previous research
	Stated Briefly Summaries of research findings for specific audiences
	Applied Research Methods Research methods for educational settings
	Tools Help for planning, gathering, analyzing, or reporting data or research