



D. Measurement

How might states and districts that lack value-added measurement systems evaluate teacher and principal performance accurately and fairly?

Signaling a growing national trend, recent high profile federal competitions such as *Race to the Top* and the *Teacher Incentive Fund* have placed a premium on using value-added student-achievement data to evaluate the performance of teachers and school leaders. Although a few states and districts scattered across the country have the assessments, data systems, and value-added models in place to facilitate this practice, the vast majority of our nation's school systems do not.¹ If a district or school has limited value-added capacity and seeks ways to accurately and fairly evaluate the performance of teachers and school leaders beyond the use of test data, it is imperative that the district use alternate assessment approaches. This paper describes two research-based approaches for evaluating the performance of teachers and principals in the absence of value-added measurement systems and standardized test data.

Performance Evaluation

Within the context of our nation's current educational landscape, one that is shaped by accountability at the district, school, and increasingly, the classroom levels, student test scores play a central role (Milanowski, Kimball, & Odden, 2005). However, education research

has failed to reach consensus over which teacher characteristics are systematically associated with students' learning gains, and it remains an open question as to whether it is even possible to judge teachers' effectiveness outside of direct observations of their teaching (Goldhaber & Anthony, 2004). According to Odden and Wallace (2007), the most powerful approaches for determining performance-based pay increases are those based on measures of teacher's instructional practice identified through a performance evaluation system. The incorporation of performance assessment data to evaluate educators' knowledge and skills is one promising approach to enhancing teacher quality (Odden, Kelley, Heneman, & Milanowski, 2001).

A quick online search confirms that many performance evaluation instruments exist for both teachers and school leaders. In the field of education, we are certainly not without data collection tools, evaluation protocols, and performance data; yet school leaders and teachers often find themselves lacking the information necessary for making crucial decisions about the effectiveness of their practice (Betebenner & Linn, 2010). The following research-based findings will help educators design and implement high-quality performance evaluation systems:

¹ For added detail, see www.dataqualitycampaign.org. While states have made significant progress developing longitudinal data systems that can follow student progress over time, only 12 states currently have data systems in place that exhibit all 10 Essential Elements of data quality.

1. For a variety of capacity, alignment, criterion validity, and inter-district mobility reasons,² states should drive the development of a performance evaluation system, allowing for important district-, school- and classroom-based contextual refinements (Cohen & Hill, 2001; Holtzapple, 2001, 2002; Odden & Wallace, 2007).
2. To accurately and fairly assess teacher and principal performance, evaluation systems should include the collection and analysis of associated artifacts.³ District and school leaders can use artifacts to ensure multiple forms of data create a rich composite of instructional and administrative practices (Bracken, Timmreck, & Church, 2001; Danielson & McGreal, 2000; Wilkerson, Manatt, Rogers, & Maughan, 2000).
3. Trained assessors must conduct the implementation of the performance evaluation system and ground it in a set of specified teaching standards and scoring rubrics to ensure valid and reliable results.⁴ Rubrics must specify low, acceptable, and high levels of performance with enough detail to make it clear what behaviors highly effective teacher or school leader must have (Danielson, 1996; Heneman, Milanowski, Kimball, & Odden, 2006; Kimball, 2002; Stronge & Tucker, 2003).
4. The performance evaluation system should directly align with professional development activities and directly link to teaching standards and district goals⁵ (Danielson & McGreal, 2000; Heneman & Milanowski, 2004; Milanowski & Kimball, 2005; Odden & Kelley, 2002).

² The research on this point contends that states should drive performance evaluation systems, allowing districts to make refinements and nuances at the local level. Many districts do not have the capacity to create a highly effective performance evaluation system on their own. Alignment is an important consideration because performance evaluation systems should link to state standards. The issue of criterion validity has to do with challenges associated with ensuring the statistical precision of the items to measure the intended domains or constructs of teacher performance. A state-based performance evaluation system helps to minimize some of the challenges associated with teacher mobility, most of which is inter-district not inter-state. A state-based approach would ensure that teachers understand and use a common set of rubrics. This will minimize the challenges of each district having its own unique and disjointed system.

³ Artifacts refer to the wide array of documentation sources that can accompany evaluations and provide data to validate teacher performance. Artifacts include lesson plans, assessments, examples of student work, and videotaped segments of teaching.

⁴ Reliability refers to the extent to which a measuring procedure yields the same results on repeated trials. Content validity refers to extent to which a measuring procedure adequately samples the domain of information, knowledge, or skill that it purports to measure.

⁵ The research contends that a performance evaluation system should be aligned with curricular standards that are guiding instruction and with the professional development system.

5. To the extent feasible, performance evaluation systems should provide interim and formative⁶ feedback to teachers and school leaders. The feedback should represent multiple perspectives (supervisor, peers, etc.) and take context and content-related pedagogy into account (Ilgen & Davis, 2000; Kimball, 2002; Kluger & DeNisi, 1996; London, 1997; London & Smither, 2002; Milanowski et al., 2001; Murphy, Elliott, Goldring, & Porter, 2006).
6. To account for varying levels of expertise, and multiple career stages of teachers, performance evaluation systems should provide differentiated processes for beginning, struggling, and experienced educators. At each career stage, the evaluation process should focus on key performance dimensions that align to district and school goals⁷ (Clotfelter, Ladd, & Vigdor, 2003; Danielson, 1996; Stronge & Tucker, 2003).
7. To ensure observation and evaluation of typical performance, it is imperative that assessors make multiple observations, at varying times of the day and school year.⁸ To minimize social pressure, multiple, trained assessors should

make the visits⁹ (Danielson, 1996; Heneman & Milanowski, 2003; Holtzapple, 2002; Ilgen & Davis, 2000; Wilkerson, Manatt, Rogers, & Maughan, 2000)

Student Learning Objectives

The second promising research-based approach to measuring the effectiveness of teachers and school leaders in states and districts lacking value-added data is the creation of student learning objectives. Student learning objectives are data-based targets of student growth that teachers set at the start of the semester or school year and strive to achieve by the end of the semester or school year. Principals approve these targets after teachers thoroughly review available student baseline data in consultation with colleagues and program support staff (Lussier & Taylor, 2007). The student learning objectives process motivates teachers to bring more science to their art and become more systematic and strategic in their instructional decisions and leads to improved teacher and student performance (Locke & Latham, 2002; Slotnik, 2008; Wallis, 2008).

⁶ The provision of one-time feedback to teachers is of limited utility. Optimally, teachers receive multiple forms of feedback throughout the year and use this information to inform their practice. This approach ensures that teachers use the feedback not merely as a summative assessment of their performance, but for formative purposes.

⁷ The provision of one-time feedback to teachers is of limited utility. Optimally, teachers receive multiple forms of feedback throughout the year and use this information to inform their practice. This approach ensures that teachers use the feedback not merely as a summative assessment of their performance, but for formative purposes.

⁸ Here “typical” refers to average or everyday performance. This is in contrast to “peak” or staged performance, which involves teachers being told in advance when once a year observation evaluations will occur.

⁹ Evaluators undergo a good deal of social pressure when just one person in the school (typically the principal) evaluates the teachers. This often results in high scores for everyone.

Researchers have analyzed the process for establishing student learning objectives for nearly a decade, and we provide the emergent research-based steps for implementing this practice below. At each step, teachers work collaboratively with a peer (such as a mentor or master teacher or an instructional coach or specialist) to accomplish each task. The school principal or program director will also collaborate with teachers, provide supporting resources, and sign off on the teachers' goals and evaluate ultimate results. The eight steps for implementing student learning objectives are as follows:

1. Conduct a needs assessment and provide rationale for goals
2. Determine specific content domains and student subgroups to target
3. Articulate learning objectives for students
4. Specify outcome assessments used to evaluate success
5. Establish student growth targets
6. Determine strategies to be used to meet objectives and targets
7. Identify professional development to support success
8. Reflect upon and appraise the process and student progress

In conclusion, while standardized test results do provide one important measure of teacher and principal effectiveness, research-based approaches for assessing educator performance exist beyond the use of value-added measurement. This paper illustrates how performance based compensation system leaders can use student learning objectives and performance evaluation systems to evaluate educator effectiveness. Fortunately, research has shown that such systems can elevate teacher and student performance and generate data that correlate to value-added student learning gains (Goldhaber, 2002; Kimball, White, Milanowski & Borman, 2004).

References

- Betebenner, D.A., & Linn, R.L. (2010). *Growth in student achievement: Issues of measurement, longitudinal data analysis and accountability*. Retrieved from <http://www.k12center.org/publications.html>
- Bracken, D.W., Timmreck, C.W., & Church, A.H. (Eds). (2001). *The handbook of multisource feedback*. San Francisco: Jossey-Bass.
- Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2003). *Teacher sorting, teacher shopping, and the assessment of teacher effectiveness*. Paper presented at the American Association of Public Policy and Management, November.
- Cohen, D.K., & Hill, H.C. (2001). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press.
- Danielson, C. (1996). *Enhancing professional practice: A framework for teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Danielson, C., & McGreal, T.L. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Goldhaber, D. (2002). The mystery of good teaching: Surveying the evidence on student achievement and teachers' characteristics. *Education Next*, 2(1), 50-55.
- Goldhaber, D., & Anthony, E. (2004). *Can teacher quality be effectively assessed?* ERIC Clearinghouse on Urban Education. New York: Teachers College, Institute for Urban and Minority Education.
- Heneman, H.G. III, & Milanowski, A.T. (2003). Continuing assessment of teacher reactions to a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(2), 171-195.
- Heneman, H.G. III, & Milanowski, A.T. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79(4), 108-125.
- Heneman, H.G. III, Milanowski, A., Kimball, S.M., & Odden, A. (2006). *Standards based teacher evaluation as a foundation for knowledge- and skill-based pay*. CPRE Policy Brief RB-45. Philadelphia, PA: Consortium for Policy Research in Education.
- Holtzapple, E. (2001). *Report on the validation of teachers evaluation system instructional domain ratings*. Cincinnati, OH: Cincinnati Public Schools.
- Holtzapple, E. (2002). *Validating a teacher evaluation system*. Paper presented at the annual meeting of the American Evaluation Association, November, Washington, DC.
- Ilgen, D.R., & Davis, C.A. (2000, July). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, 49(3), 550-565.
- Kimball, S.M. (2002). Analysis of feedback, enabling conditions and fairness perceptions of teachers in three school districts with new standards-based evaluation systems. *Journal for Personnel Evaluation in Education*, 16, 241-268.
- Kimball, S.M., White, B., Milanowski, A.T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student assessment results in Washoe County. *Peabody Journal of Education*, 79(4), 54-78.
- Kluger, A.N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.

References continued

- Locke, E.A., & Latham, G.P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist* (September 2002), 706.
- London, M. (1997). *Job feedback: Giving, seeking, and using feedback for performance improvement*. Mahwah, NJ: Lawrence Erlbaum Associates.
- London, M., & Smither, J.W. (2002, Spring). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review*, 12(1), 81-100.
- Lussier, D., & Taylor, J. (2007). *Student learning objectives*. AISD Strategic Compensation Program. Retrieved from <http://www.austinisd.org/inside/initiatives/compensation/slos.phtml>
- Milanowski, A.T., & Kimball, S.M. (2005). *The relationship between teacher expertise and student achievement: A synthesis of three years of data*. Paper presented at the American Educational Research Association Meeting. Montreal, Quebec, Canada.
- Milanowski, A.T., Kimball, S.M., & Odden, A. (2005). Teacher accountability measures and links to learning. In L. Stiefel, A.E. Schwartz, R. Rubenstein, & J. Zabel (Eds.). *Measuring school performance and efficiency: Implications for practice and research*, 137-161. Larchmont, NY: Eye on Education.
- Milanowski, A., Kellor, E., Odden, A., Heneman, H.G. III, White, B., Allen, J., et al. (2001, July). *Final report on the evaluation of the 2000-2001 implementation of the Cincinnati federation of teachers/Cincinnati public schools teacher evaluation system* (Working Paper No. TC-01-3). Madison: University of Wisconsin, Wisconsin Center for Education Research, Consortium for Policy Research in Education.
- Murphy, J., Elliott, S., Goldring, E., & Porter, A. (2006). *Learning centered leadership: A conceptual foundation*. Prepared for the Wallace Foundation.
- Odden, A., & Kelley, C. (2002). *Paying teachers for what they know and can do: New and smarter compensation strategies to improve student achievement*. Thousand Oaks, CA: Corwin Press.
- Odden, A., & Wallace, M.J. Jr. (2007). *How to create world class teacher compensation*. St. Paul, MN: Freeload Press.
- Slotnik, W. (2008). *Student learning objectives. Teacher compensation brief*. Community Training and Assistance Center. Boston, MA.
- Stronge, J.H., & Tucker, P.D. (2003). *Handbook on teacher evaluation: Assessing and improving performance*. Larchmont, NY: Eye on Education, Inc.
- Wallis, C. (2008). How to make great teachers. *Time*, 171(8), 28-34.
- Wilkerson, D.J., Manatt, R.P., Rogers, M.A., & Maughan, R. (2000, June). Validation of student, principal, and self-ratings in 360 degree feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 179-192.

This synthesis of key research studies was written by:
Patrick J. Schuermann, Vanderbilt University

The Center for Educator Compensation Reform (CECR) was awarded to Westat — in partnership with Learning Point Associates, Synergy Enterprises Inc., Vanderbilt University, and the University of Wisconsin — by the U.S. Department of Education in October 2006.

The primary purpose of CECR is to support Teacher Incentive Fund (TIF) grantees in their implementation efforts through provision of sustained technical assistance and development and dissemination of timely resources. CECR also is charged with raising national awareness of alternative and effective strategies for educator compensation through a newsletter, a Web-based clearinghouse, and other outreach activities.

This work was originally produced in whole or in part by the CECR with funds from the U.S. Department of Education under contract number ED-06-CO-0110. The content does not necessarily reflect the position or policy of CECR or the Department of Education, nor does mention or visual representation of trade names, commercial products, or organizations imply endorsement by CECR or the federal government.



Center for
Educator Compensation
Reform

Allison Henderson, Director
Phone: 888-202-1513
E-mail: cecr@westat.com