

**DESIGNING PRINCIPAL EVALUATION SYSTEMS:
RESEARCH TO GUIDE DECISION-MAKING**
An Executive Summary of Current Research

Compiled, Analyzed, and Written by
Matthew Clifford, Ph.D., American Institutes for Research
Steven Ross, Ph.D., Center for Research and Reform in Education,
Johns Hopkins University

Prepared in Collaboration with the
National Association of Elementary School Principals

July 14, 2011



AMERICAN
INSTITUTES
FOR RESEARCH®

This document is a summary of current research on the condition of principal evaluation systems and the relationship between principal evaluation and principals' capacity to lead learning communities. It is not intended to present new research, but rather to create a foundation for further exploration and analysis. It was developed by Matthew Clifford, Ph.D., American Institutes for Research, and Steven Ross, Ph.D., Center for Research and Reform in Education, Johns Hopkins University, in collaboration with the National Association of Elementary School Principals (NAESP).

Copyright © 2011 by American Institutes for Research (AIR). All rights reserved. No part of this publication may be reproduced or transmitted in any form, by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without attribution or the permission of AIR.

Introduction

High-quality school leadership is essential to achieving our national goals of making dramatic changes in the lowest performing schools and improving the educational opportunities for all students. Increasingly, principals and assistant principals are called upon to engage deeply with teachers in designing classroom and school learning environments that are academically rigorous, safe, and supportive of students' social and emotional learning.

Performance evaluation can be a powerful way to support the continuous growth and development of principals as instructional leaders. Prompted by Race to the Top (RTTT) and other initiatives, many states are now redesigning principal evaluation systems as one means of providing support for principals' work. Ideally, principal evaluation systems reflect new goals and priorities for principals and support improved school leadership.

The scope of principal evaluation systems improvement efforts is, in many ways, far reaching, and the work is complex. American Institutes for Research and Johns Hopkins University staff are conducting a review of literature on principal evaluation to ascertain the status of research on the topic in order to inform state, district, and other organizations' efforts to improve principal evaluation systems. This executive brief summarizes the research review and raises questions that may be useful to the field.

Evidence that Principals Matter

A focus on principal evaluation makes sense because principals' work has a strong, measurable effect on student achievement, teaching quality, and schools. Research evidence confirms what policymakers, practitioners, and others have long understood: School principals have a strong effect on student achievement, in creating conditions and cultures that lead to better teaching and learning, and on shaping the long-term impact of school improvements.

Multiple research studies indicate a strong association between principals' work and student achievement. Several recent meta-analyses conclude that school leadership is second only to classroom instruction among school-related factors contributing to student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Murphy & Datnow, 2003; Supovitz & Pogliano, 2001; Marzano, Waters, & McNulty, 2005). Some researchers posit that the positive impact of leadership is even greater in low-performing schools (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Branch, Hanushek, & Rivkin, 2009; Hallinger & Heck, 1998).

In addition, principals' work also has been shown to impact school performance, teacher quality, and policy implementation. These leadership functions can indirectly influence student learning (Hallinger & Heck, 1998). For example, some studies of schools and teachers indicate that principals:

- Are a determining factor in teachers' decisions to join and remain in a school (Boyd, Grossman, Ing, Lankford, Loeb, & Wykoff, 2010; Ladd, 2009);
- Set and maintain school reform agendas (Gamoran et al., 2003);
- Build organizational culture and trust in school buildings (Bryk, Schneider, Greenberg, & Kochanek, 2002);
- Affect schools' abilities to administer programs and influence change through the allocation of human, financial, and other resources (Leithwood & Montgomery, 1982); and
- Help interpret district and state policies, thereby influencing if and how policies are enacted (Halverson & Clifford, 2006; Zmach, 2006).

Evidence of the Need for Better Principal Evaluation

Principal evaluation has long held promise for improving principals' practice, building their capacity, holding them accountable for student progress, and ensuring that they have an overall positive impact on students and schools. According to experts, well-structured principal performance assessment procedures can reinforce and strengthen leadership practices when evaluations provide timely, trustworthy feedback on performance and offer districts mechanisms for monitoring principal progress (Goldring et al., 2009; Halverson, Kelley, & Kimball, 2004; Kimball, Milanowski, & McKinney, 2007). Feedback on practice is particularly important because principals report that they have few sources of trusted feedback on their practice and commonly feel isolated from colleagues due to the rigors of their position (Friedman, 2002).

Given the connection between effective school leadership and student achievement, evaluating and developing school principals is increasingly recognized as a key strategy for improving schools, increasing student achievement, and narrowing student achievement gaps. In 2001, the No Child Left Behind Act required states to adopt new methods for evaluating principals that included high-stakes, summative measures of student performance while also addressing principal preservice and inservice quality improvements. As part of the Obama administration's RTTT and school improvement agenda, teacher and principal evaluation are emphasized due to the requirement that states and local districts determine "effective" and "highly effective" teachers and school leaders.

While principal evaluation holds great potential, a relatively small number of studies on principal evaluation practices are available, and those suggest that improvements are long overdue. The studies raise questions about the consistency, fairness, effectiveness, accountability, and value of current principal evaluation practice. Research on principal evaluation also suggests that:

- Principals view performance evaluation as perfunctory, having limited value for feedback, professional development, or accountability to school improvement (Portin, Feldman, & Knapp, 2006);

- Principal evaluations are inconsistently administered; therefore, performance is inconsistently measured (Thomas, Holdaway, & Ward, 2000);
- Performance evaluations may not align with existing state or national professional standards for practice (Heck & Marcoulides, 1996) or standards for personnel evaluation (Goldring et al., 2009); and
- Few widely available principal evaluation instruments display psychometric rigor or make testing public so that validity and reliability can be examined (Condon & Clifford, 2010; Goldring et al., 2009; Heck & Marcoulides, 1996).

Framing Principal Evaluation: Policies and Standards that Guide Design

While research highlights the shortcomings of current principal evaluation approaches, less literature is available on the effects of evaluation on the growth and development of leaders and the improvement of schools and students. Such research is important for providing states, districts, and others with more certain guidance on principal evaluation design. Catano and Stronge (2007) observe that the majority of educator evaluation research focuses on the evaluation of teachers, not principals. While some lessons can be taken from research on teacher evaluation systems, the work and responsibilities of principals differ greatly than that of teachers and should be evaluated differently.

Although additional research is needed on principal evaluation design and effects, current standards and policies frame the work of states and districts as they design—or make recommendations for the design of—principal evaluation. An extensive review of standards, policies, and related research are needed to provide clear guidance to the field on evaluation design. The standards and policies highlighted here are research based and reflect current expert opinion.

National professional standards of leadership practice have been developed through extensive effort by multiple organizations; they can serve as a starting point for principal evaluation design. These standards have been integrated into state and district principal professional standards and into many preservice, inservice, and other programs. Standards for professional practice represent a major step toward developing principal evaluations (see Stufflebeam & Nevo, 1993). A comprehensive review and categorization of professional standards has not recently been conducted, however, nor has a review of standards against criteria for personnel evaluation. That said, a few standards for practice have been adopted by states and large school districts. These may serve as a starting point for additional review and evaluation design.

- **Interstate School Leadership Licensure Consortium (ISLLC) standards** and indicators have been produced through extensive review of principal and school effectiveness literature (Council of Chief State School Officers, 2008). These have been adopted by a majority of states for performance evaluation and preparation purposes (Hale & Moorman, 2003; Anthes, 2005).

- **National Board of Professional Teaching Standards: Standards for Principals** have been developed to guide principal development through an extensive review of research literature and expert input. They are intended to guide principal development as instructional leaders and underpin the National Board’s master principal assessment system.
- **The National Association of Elementary School Principals’ Leading Learning Communities: Standards for What Principals Should Know and Be Able to Do** focus on the role of principals as instructional leaders and participants in learning communities within schools that create conditions for continuously improving student learning.

In addition to these nationally recognized, research-based standards for school leaders, other individuals and organizations have extensively reviewed research to create standards for leadership practice that have informed state and district evaluation systems.

In preparation for the RTTT competition and other related initiatives, many states and districts have changed, or will soon change, principal evaluation policies and have begun reforming evaluation systems. Research from the field of human resources and educational human capital management can provide states and districts with guidance on evaluation design, and the Joint Committee on Standards for Educational Evaluation’s *Personnel Evaluation Standards* (2010) also may be informative. While the below list of evaluation systems features is by no means comprehensive, it may provide a starting point for policymakers, evaluation designers, and others. Our review suggests that principal evaluation systems should:

- **Be designed with the direct involvement of principals and other constituents:** Principal evaluation design should engage school-level leaders and their representatives in the design process, so that their experience and insight is reflected in both the content and process of evaluation. Engaging leaders in the process builds trust and credibility for new evaluation systems and ensures that the evaluation process is feasible and useful to administrators.
- **Be educative:** A principal evaluation system should provide useful, valuable, and trustworthy data and feedback to principals in order to advance principals’ abilities to be more effective leaders within their schools and communities.
- **Be connected to district- and state-level systems:** Principal evaluation should be considered one component of a broader leadership human capital management system. Data arising from performance evaluations can be used to design professional development and induction systems, shape hiring procedures, improve working conditions, develop incentives, and inform other human resource processes that support leaders.
- **Be rigorous, fair, and equitable:** The content, instruments, and administration of principal evaluation systems should be legal and ethical; allow for a thorough examination of principal practice; and be valid, reliable, and accurate.

- **Include multiple rating categories to differentiate performance:** Principal evaluation should be based on transparent expectations for performance that are attuned to school leaders' work by taking into account, possibly, differences between secondary, middle, and elementary school leadership approaches. It should clearly differentiate levels of performance.
- **Gather evidence of performance through multiple measures of practice:** Evaluations should use multiple measures to provide a holistic view of principal performance. These measures might include, but are not limited to, 360-degree surveys of staff, observations, school climate or other surveys, and weighted summative measures of school and student performance.
- **Communicate results to principals consistently and with transparency:** Principal evaluations are powerful to the extent that feedback can be used by principals to improve their work in schools and by district staff to make personnel decisions. Feedback should include all data from evaluations and should be clear, pointed, and actionable.
- **Include training, support, and evaluation of principal evaluators:** New evaluation systems should be administered with consistency and fidelity, which requires that evaluators are trained, monitored, and supported.

Potential Questions for Consideration

The review of literature points both to the importance of principals in attaining school and student performance goals and to the need for improved principal evaluation to continuously develop these leaders' abilities to inspire others and create conditions for high-quality teaching and improved student learning. This review also raises questions that might be considered by researchers, policymakers, and others concerned with principal evaluation design. Answering these and other questions may help states, districts, and others to design comprehensive, fair, and objective evaluation instruments, processes, and systems.

Questions for Policymakers and Evaluation Designers

What are the purposes of the evaluation systems? The research review notes that principal evaluation can serve many purposes. A clear understanding of these purposes would likely be informative to evaluation design.

What should be assessed? While a holistic assessment of principal performance is desired, the scope of principals' work raises the need to prioritize what can and should be assessed.

How, if at all, should evaluation be differentiated for elementary, middle-level, and secondary principals? Research suggests elementary and secondary principals' work differs in many ways. Evaluation designs may need to take these differences into account.

How, if at all, should evaluation be differentiated by principals' level of expertise? Research indicates that principals, like other professionals, develop throughout their careers. Guidance

might be provided on how, if at all, evaluation should be differentiated for these developmental levels.

What measures should be used? Multiple measures, such as observations or 360-degree surveys, have been recommended for use in evaluating principals, but how should evaluation system designers determine which measures to use? What standards should be used to determine which measures are unbiased, fair, and reliable?

How can evaluation be made feasible, fair, valid, and routine? Standards for principal and personnel evaluation highlight the importance of feasible, valid, and routine evaluations, but research provides little guidance on ways evaluation can be structured to meet these criteria.

How can evidence of performance be integrated? While multiple measures of principal performance are recommended, this evidence will likely need to be weighted and represented in ways that reflect leadership standards and priorities.

Who should evaluate principals? Local evaluators will be responsible for ensuring that principal evaluation is fair, valid, and reliable, and that feedback is actionable. While principals' direct supervisors are a logical choice as evaluators, not all districts will likely have the same capacity to conduct principal evaluations. For example, some principals also are superintendents.

What must principal evaluators know and do, and how can their work be supported?

Evaluating principals is a complex practice requiring expertise and support, but little research or guidance is available on what principal evaluators must know and do to be effective.

How can principal evaluation data best be collected and used? Improved principal evaluation holds potential for informing other aspects of district and state human resources systems, and the data collected can be used to monitor the quality of a state or district's principal evaluation system. But good data will need to be gathered, maintained, and used in order to meet this potential.

What is the best process for states and districts to use when redesigning principal evaluation systems? Standards and policies recommend that evaluation systems be redesigned with extensive input from principals and others. Further, the task of principal evaluation design is highly complex. Research and other literature provide little guidance on how to engage constituents or move the design process forward.

What resources, assistance, or knowledge is beneficial to support principal evaluation system design, and who is best positioned to supply that knowledge? Little information is available on the types of information needed by states and districts as they redesign principal evaluation systems.

Questions for Researchers

What types of information and feedback do principals find valuable and need? Existing research provides little guidance on the types and forms of information that are helpful to principals and other constituents.

How valid, reliable, and feasible are common measures of principal effectiveness? Few comprehensive reviews of the psychometric properties of existing measures of principal effectiveness have been conducted. These reviews can be very informative to policymakers and administrators seeking to adopt off-the-shelf instruments.

What are the effects of principal evaluation systems? Few quasi-experimental or experimental studies of principal evaluation systems have been conducted to determine the effects, if any, on principals' development, teaching quality, school effectiveness, or student achievement.

What conditions must exist in states or districts for principal evaluation to achieve scale and be effective? The review located no studies that examine school, district, or state conditions amenable to scaled or effective principal evaluation systems.

References

- Anthes, K. (2005). *Leader standards*. Denver, CO: Education Commission of the States (ECS).
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2010). The influence of school administrators on teacher retention decisions. *American Education Research Journal, 48* (2), 303–333.
- Branch, G., Hanushek, E., & Rivkin, S. (2009). *Estimating principal effectiveness*. (CALDER Working Paper 32). Washington, DC: Urban Institute.
- Bryk, A., Schneider, B., Greenberg, S., & Kochanek, J. (2002). *Trust in schools: A core resource for improvement*. Washington, DC: American Sociological Association.
- Catano, N., & Stronge, J. (2007). What do we expect of school principals? Congruence between principal evaluation and performance standards. *International Journal of Leadership in Education, 10*(4), 379–399.
- Condon, C., & Clifford, M. (2010). *Measuring principal performance: How rigorous are commonly used principal performance assessment instruments?* Naperville, IL: Learning Point Associates.
- Council of Chief State School Officers. (2008). *Educational leadership policy standards: ISLLC 2008*. Washington, DC: Author.
- Friedman, I. (2002). Burnout in school principals: Role related antecedents. *Social Psychology of Education, 5*(3), 229–251.
- Gamoran, A., Anderson, C. W., Quiroz, P. A., Secada, W. G., Williams, T., & Ashman, S. (2003). *Transforming teaching in math and science: How districts and schools can support change*. New York: Teachers College Press.
- Goldring, E., Carvens, X., Murphy, J., Porter, A., Elliott, S., & Carson, B. (2009). The evaluation of principals: What and how do states and urban districts assess leadership? *Elementary School Journal, 110*(1), 19–39.
- Hale, E. & Moorman, H. (2003). *Preparing school leaders: A national perspective on policy and program innovations*. Washington, D.C.: Institute for Educational Leadership.
- Hallinger, P., & Heck, R. H. (1998). Exploring the principal's contribution to school effectiveness: 1980–1995. *School Effectiveness and School Improvement, 9*, 157–191.
- Halverson, R., & Clifford, M. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Education Administration Quarterly, 42*(4), 578–619

- Halverson, R., Kelley, C., & Kimball, S. (2004). Implementing teacher evaluation systems: Principals make sense of complex artifacts to shape local instructional practice. In W. K. Hoy & C. G. Miskel (Eds.), *Administration, Policy and Reform: Research and Measurement Research and Theory in Educational Administration*. Greenwich, CT: Information Age.
- Heck, R. H., & Marcoulides, G. A. (1996). Principal assessment: Conceptual problem, methodological problem, or both? *Peabody Journal of Education*, 68(1), 124–144.
- Joint Committee on Standards for Educational Evaluation. (2010). *Personnel Evaluation Standards*. Iowa City, IA: Author. Retrieved from <http://www.jcsee.org/personnel-evaluation-standards>
- Kimball, S., Milanowski, T., & McKinney, S. (2007, April). *Implementation of standards-based principal evaluation in one school district: First year results from randomized trial*. Paper presented at the annual conference of the American Educational Research Association, Chicago. Retrieved from <http://cpre.wceruw.org/publications/KimballMilanowskiMcKinney.pdf>
- Ladd, H. (2009). *Teachers' perceptions of their working conditions: How predictive of policy-relevant outcomes?* (CALDER Working Paper 33). Washington, DC: Urban Institute.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. St. Paul, MN: University of Minnesota, Center for Applied Research and Educational Improvement & Toronto, Canada: Ontario Institute for Studies in Education.
- Leithwood, K., & Montgomery, D. (1982). The role of the elementary school principal in program improvement. *Review of Educational Research*, 52(3), 309–339.
- Marzano, R. J., Waters, T., & McNulty, B. A. (2005). *School leadership that works: From research to results*. Alexandria, VA: ASCD.
- Murphy, J., & Datnow, A. (2003). *Leadership lessons from comprehensive school reform*. San Francisco: Corwin Press.
- Portin, B. S., Feldman, S., & Knapp, M. S. (2006). *Purposes, uses, and practices of leadership assessment in education*. New York: The Wallace Foundation.
- Stufflebeam, D., & Nevo, D. (1993). Principal evaluation: New directions for improvement. *Peabody Journal of Education*, 68(2), 24–46.
- Supovitz, J., & Poglinco, S. (2001). *Instructional leadership in a standards-based reform*. Philadelphia: Consortium for Policy Research in Education.

Thomas, D., Holdaway, E., & Ward, K. (2000). Policies and practices involved in the evaluation of school principals. *Journal of Personnel Evaluation in Education*, 14(3), 215–240.

Zmach, C. (2006). *Navigating the complexities of legislation: How elementary school principals interpret and implement Florida's third grade retention policy*. Jacksonville, FL: University of Florida.