

7-1-2015

## The Role of Performance Management in Good Governance and Its Application in Public Education

Carolyn J. Heinrich  
*University of Texas at Austin*

Follow this and additional works at: [https://research.upjohn.org/empl\\_research](https://research.upjohn.org/empl_research)



Part of the [Education Economics Commons](#)

---

### Citation

Heinrich, Carolyn, J. 2015. "The Role of Performance Management in Good Governance and Its Application in Public Education." *Employment Research* 22(3): 4-6. [https://doi.org/10.17848/1075-8445.22\(3\)-2](https://doi.org/10.17848/1075-8445.22(3)-2)

This title is brought to you by the Upjohn Institute. For more information, please contact [repository@upjohn.org](mailto:repository@upjohn.org).

Carolyn J. Heinrich

# The Role of Performance Management in Good Governance and Its Application in Public Education

*This article draws from the author's chapter in The Political Economy of Good Governance (Asefa and Huang, eds.), which was recently published by the Upjohn Institute. To order the book, visit [www.upjohn.org/up\\_press](http://www.upjohn.org/up_press), or see p. 7 for more details.*

**G**overnance—laws, rules, judicial decisions, and administrative practices that prescribe and enable the provision of publicly supported goods and services—determines government performance (Lynn, Heinrich, and Hill 2001). The role of performance management, in turn, is to shape how public sector agencies, programs, and activities are organized and managed to achieve public purposes and desired outcomes.

The origins of performance management lie in a basic agency-theory framework, where an owner hires managers and workers to generate profits (with the owner or manager acting as principal, and the workers as agents). The principal's main objective is to design a contract that aligns principal and agent incentives and achieves the principal's production objectives. This is made challenging, however, by the fact that these relationships are frequently typified by conflicts in goals and values, as well as privately held information or information asymmetries.

It is here that a role for performance management enters in, in monitoring worker actions, outputs, and outcomes, and in developing an incentive scheme that aligns principal and agent interests—essentially, a contractual relationship with performance expectations and

credible provisions for enforcing it. However, even in a simple production system—where organizational goals and production tasks are known, a linear relationship exists between efforts and outputs, and there are relatively few variables for managers to control—an enforceable contract is difficult to achieve.

One well-known problem is adverse selection, where employees' true motivations or capabilities for producing a desired outcome are unknown. The second is moral hazard and unobservability, in which employees'

---

### **Effective performance management demands clarity of goals and their translation into empirical measures that adequately characterize our intended outcomes.**

---

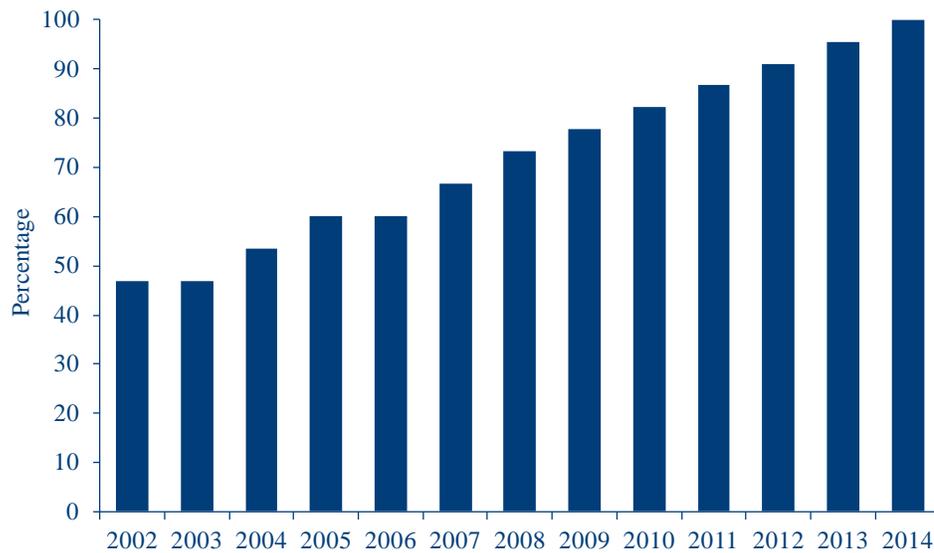
efforts or actions are not observable or readily measured, creating conditions that encourage shirking or distorted results. Recent headlines reporting cheating scandals in K–12 schools—under pressure to meet performance targets on standardized tests set by the No Child Left Behind (NCLB) Act—are just one example of how these problems can undermine performance management efforts (Rich 2013).

Yet many contracts and performance management systems still incorporate basic linear (or “straight-line”) incentive schemes, largely because of their perceived simplicity and the significant

costs of establishing a more intricate contract or system of incentives. A straight-line approach typically defines a required (linear) rate of performance improvement from an initial score or target and may also specify an ending value corresponding to a maximum performance level, such as NCLB's goal of 100 percent proficiency in reading and mathematics for public school students (see Figure 1). NCLB also provides an example of an important shortcoming of straight-line models for establishing performance expectations: they are seldom constructed using empirical data that would generate realistic expectations for performance (Koretz and Hamilton 2006). In fact, Secretary of Education Arne Duncan acknowledged that the performance management system under NCLB evolved “from an instrument of reform into a barrier to reform” (U.S. Department of Education 2013, p. 1).

From the start, the application of agency theory to the design of performance management systems in the public sector has been complex. First, just *who* is the principal in a given governance setting? Governance in the public sector is multilayered and dynamic. In addition, consensus or clarity on goals is often lacking among citizens, and sometimes in originating legislation as well. The public sector is also distinctive in that its primary work typically involves complex, nonmanual work, characterized by multilevel interactions and public-private sector partnerships. Nonstandardized outputs make the accurate measurement of performance and construction of performance benchmarks more challenging and more costly. Finally, the public sector is also distinct from the private sector in the extent to which political influences may be brought to bear at many different levels. Goals and priorities can change swiftly, and entire agencies or authority structures can be reorganized, as well as the foci of primary work. There is great potential for unintended consequences as performance management and the use of performance-based contracts expand into uncharted public-sector territory (Koning and Heinrich 2013).

**Figure 2.1 Annual Expectations Set by the No Child Left Behind Act for Increased Performance among Students in Grades K–8 Tested in Math, 2002–2014 (% that must be met of students testing at the “proficient” level)**



NOTE: The figure shows the expected percentage of students each year that should achieve a rating of “proficient” in testing for math under the legislation. The bar for 2002, the first year the law was in effect, shows the actual percentage of students who tested as proficient in math that year, and no improvement was required for the first year following that. But thereafter, increasing percentages of proficiency were set for each year, culminating in 2014, when 100 percent of kindergarteners through eighth graders were expected to be proficient in math. (The exception was 2006, when the expectations were not raised from 2005.)

SOURCE: Author’s compilation.

### Applying Performance Management Bluntly in Public Education

We spend close to \$600 billion annually on our public elementary and secondary school system, and the public is demanding greater accountability and results. Furthermore, public education today is characterized by elaborate governing structures with deeply layered and overlapping levels of decision making, widely varying views on appropriate means and ends for improving education, an increasingly complex technology with diverse outputs (which we subject to standardized measures of outcomes), and political influences that interject at many levels. Could the use of performance management potentially bring some clarity and coherence to K–12 education governance?

We have proceeded full speed ahead with regimes for performance management and accountability in education that include strong incentives and high-stakes consequences for

many stakeholders. NCLB marked the beginning of an assertive federal role in directing state and local practices to meet student performance standards. The federal government holds states, districts, and schools accountable for a comprehensive set of standards, including annual academic progress, teacher quality, and achievement gaps, and for developing assessments of student performance relative to those standards. NCLB defines educational success primarily based on standardized tests of students’ performance, and current funding and accountability systems presume “same-age cohorts of students proceeding in lockstep” (Wilson 2013, p. 96). Consistent with the origins of performance management, Darling-Hammond (2002, p. 6) describes how our test-based accountability system reflects a “factory-model approach” to education, in which schools are organized “to process large batches of students in assembly-line fashion rather than to ensure that students are well-known

by their teachers and treated as serious learners.”

Recently, recognition of the limitations of proficiency measures under NCLB has propelled alternative approaches to measuring educational performance, particularly value-added measures. A basic value-added model compares the individual growth of a group of students (e.g., in a given classroom or school) to average growth of the population of interest (e.g., growth among all students in the state). Some value-added models are also constructed to account for factors outside the control of schools in estimating growth in student achievement over time. Although these are (arguably) better measures of performance than proficiency levels, should society be ratcheting up the stakes that it attaches to them, as we have recently seen in some large, urban school districts?

One of the most controversial recent developments in performance management in education has been the high-profile, public dissemination of value-added measures of teacher performance in large school districts, including in Los Angeles and New York. Calculated by third parties (outside the district), the value-added measures associated with specific teachers were published in the *Los Angeles Times* and by the New York City Department of Education.<sup>1</sup> The objective was to get the performance information directly to citizen stakeholders, who could use this information and their political power to drive public-sector performance improvements.

However, in New York City, the margin of error in value-added measures was so wide that the average confidence interval around each rating spanned 35 percentiles in math and 53 percentiles in English, the city said. Some teachers were judged on as few as 10 students. In publishing the Los Angeles numbers, the *L.A. Times* acknowledged that value-added measures “do not capture everything about a teacher or school’s performance” (see Note 1). A study by Mathematica Policy Research (Schochet and Chiang 2010) finds that the error rate for value-added scores (based on three years of data) was 25 percent. Therefore, a three-year model would rate one out

of every four teachers incorrectly, and with only one year of data, the error rate jumped to 35 percent.

### Lessons for Improving the Effectiveness of Performance Management

What have we learned about the role of performance management in contributing to good governance and improving government outcomes?

- The effective use of performance management demands clarity of goals and their translation into empirical measures that accurately and adequately characterize our intended outcomes.

Where we fail on either of these components, the performance management system may risk doing more harm than good. In many cases, the data available simply are not up to the task.

In light of these limitations, and recognizing that performance management often grapples with multiple goals and complex production, we may be better off with multidimensional measures of performance to guide our work. A number of school districts and states are now developing these types of multipurpose, multiple-indicator performance management systems for K–12 education (New York City Department of Education 2014). A potential trade-off, of course, is that a more intricate or complicated system and set of incentives would likely place a greater demand on public capacities for managing such a system.

- Caution should be exercised in attaching high stakes to performance results, given the known challenges and imperfections of our performance measures.

The awarding of performance bonuses, “naming and shaming” (as in the publication of teacher value-added), termination of contracts, or retractions of program funding would best be backed or verified by multiple sources of quantitative and qualitative evidence before going forward. A counterargument frequently offered against eliminating high stakes altogether is that the performance management and incentive systems will lose their “teeth” and purpose. Evidence to date,

however, suggests that individuals and organizations are highly responsive to performance standards, even when the rewards are minimal, such as peer recognition (Bevan and Hood 2006; Heinrich 2007).

- Performance management systems are likely to be more effective tools of governance if we focus more on their use for *diagnostic* purposes.

That is, resources and rewards should follow their *effective use* in improving government and program outcomes, rather than for hitting performance targets. In the public education example, schools or teachers would be rewarded for using information on students’ performance to help increase their learning, ideally measured in terms of their individual growth that is not based solely on test score levels or gains. This would be a more appropriate outcome to report publicly (for the sake of transparency), and, if measured sufficiently, would also reward the right types of efforts to increase performance, that is, not success in increasing test-taking skills but rather effective use of performance information to help students succeed academically.

### Note

1. See <http://projects.latimes.com/value-added/> and <http://www.schoolbook.org/2012/02/24/teacher-data-reports-are-released/> (accessed June 10, 2015).

### References

Bevan, Gwyn, and Christopher Hood. 2006. “What’s Measured Is What Matters: Targets and Gaming in the English Public Health Care System.” *Public Administration* 84(3): 517–538.

Darling-Hammond, Linda. 2002. “Standards, Assessments, and Educational Policy: In Pursuit of Genuine Accountability.” The eighth annual William H. Angoff Memorial Lecture, presented at the Educational Testing Service, Princeton, NJ, October 30.

Heinrich, Carolyn J. 2007. “False or Fitting Recognition? The Use of High Performance Bonuses in Motivating Organizational Achievements.” *Journal of Policy Analysis and Management* 26(2): 281–304.

Koning, Pierre, and Carolyn J. Heinrich. 2013. “Cream-Skimming, Parking, and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts.” *Journal of Policy Analysis and Management* 32(3): 461–483.

Koretz, Daniel M., and Laura S. Hamilton. 2006. “Testing for Accountability in K–12.” In *Educational Measurement*, Robert L. Brennan, ed. 4th ed. American Council on Education–Praeger Series on Higher Education. Westport, CT: Praeger, pp. 531–578.

Lynn, Laurence E. Jr., Carolyn J. Heinrich, and Carolyn J. Hill. 2001. *Improving Governance: A New Logic for Empirical Research*. Washington, DC: Georgetown University Press.

New York City Department of Education. 2014. *Performance and Accountability*. New York: New York City Department of Education. <http://schools.nyc.gov/Accountability/default.htm> (accessed October 16, 2014).

Rich, Motoko. 2013. “Scandal in Atlanta Reignites Debate over Tests’ Role.” *New York Times*, April 3, A:13. [http://www.nytimes.com/2013/04/03/education/atlanta-cheating-scandal-reignites-testing-debate.html?\\_r=0](http://www.nytimes.com/2013/04/03/education/atlanta-cheating-scandal-reignites-testing-debate.html?_r=0) (accessed September 17, 2014).

Schochet, Peter Z., and Hanley S. Chiang. 2010. *Error Rates in Measuring Teacher and School Performance Based on Student Test Score Gains*. NCEE Report No. 2010-4004. Washington, DC: U.S. Department of Education, National Center for Education Evaluation and Regional Assistance.

U.S. Department of Education. 2013. *No Child Left Behind: Early Lessons from State Flexibility Waivers*. Testimony of Secretary of Education Arne Duncan to the U.S. Senate Committee on Health, Education, Labor, and Pensions, February 7. 113th Cong. Washington, DC: U.S. Department of Education. <http://www.ed.gov/news/speeches/no-child-left-behind-early-lessons-state-flexibility-waivers> (accessed October 16, 2014).

Wilson, Woodrow. 1887. “The Study of Administration.” *Political Science Quarterly* 2(2): 197–222.

*Carolyn J. Heinrich is the Sid Richardson Professor of Public Affairs and affiliated professor of economics and the director of the Center for Health and Social Policy at the University of Texas at Austin.*