

# **Professional Accountability as a Tool for Continuous Improvement in the Age of ESSA**

**Brian P. Gill, Mathematica Policy Research**

**Spring 2017 AAFP conference**

## **Accountability comes in multiple forms**

Since the passage of the No Child Left Behind Act (NCLB), American policymakers have understood “accountability” in K-12 education to mean something very specific: formal consequences for schools and/or educators based on student outcomes (typically test scores). Outcome-based accountability in the form of high-stakes testing has been the primary policy lever used by policymakers over the last two decades to promote improved school performance.

In other sectors outside of K-12 schooling, however, accountability takes many different forms; attaching formal consequences to outcomes is only one of these forms. Elected officials are accountable through the ballot box. Doctors and lawyers are accountable through professional standards. Consumer-serving businesses are accountable through the market. These varied forms of accountability reflect the fact that, psychologically, accountability can be created in several different ways. The behavioral science literature, based largely on randomized experimental studies in social psychology and behavioral economics, has identified four discrete behavioral accountability mechanisms (Lerner & Tetlock 1999):

1. Evaluation: the expectation that consequences depend on one’s performance.
2. Identifiability: The expectation that an action or outcome is attributable to oneself.
3. Reason-giving: The expectation that one needs to explain or rationalize one’s actions.
4. Mere presence of another: The simple awareness that someone is watching.

High-stakes testing, like other forms of outcome-based accountability, relies on evaluation and identifiability. If accountability policymakers understand accountability exclusively in terms of consequences for outcomes, they will overlook other forms of accountability that invoke reason-giving and the mere presence of another. In a previous paper (Gill, Lerner, and Meosky 2016), we explored the implications of behavioral science findings for accountability in schools, describing three other forms of accountability—rule-based, market-based, and professional accountability—that can be used alongside outcome-based accountability, collectively invoking all four of the behavioral accountability mechanisms. The table below, reprinted from that paper, provides examples of specific policies and practices for each of the accountability types and corresponding behavioral/psychological mechanisms.

**Accountability types and psychological accountability mechanisms with applications in K-12 schooling**

		Accountability types			
		Outcome-based	Rule-based	Market-based	Professional
Psychological accountability mechanisms	Mere presence of another				Classroom windows
	Identifiability	Public reporting of schoolwide test results	Minimum certification requirements	School choice	Peer observation
	Reason-giving			Charter-school enrollment fairs	Instructional coaching
	Evaluation	Value-added incorporated in teacher evaluations	Formal observation by principal		Peer review; advanced certification

We concluded that *professional accountability* may be unique in its ability to (potentially) employ all four psychological mechanisms and in coupling accountability with support for improvement. It has also historically been under-used in K-12 schooling and therefore merits

more attention as an accountability type that could be used in conjunction with other forms of accountability.

The current paper describes how policymakers and educators might employ professional accountability to address professional practice, using examples from schools that are already doing so. Professional accountability systems can provide not only more robust incentives but also support to help educators get better at their work. Used in conjunction with outcome-based and other forms of accountability, professional accountability creates the potential for a cycle of continuous improvement in schools.

Consideration of a wider range of accountability types is particularly relevant now that the Every Student Succeeds Act (ESSA) has replaced NCLB, giving states greater discretion in the design of their accountability systems. There is a new opportunity for policymakers and educators to employ accountability tools that more fully activate the psychological accountability mechanisms that have been identified in the behavioral science literature. The public discussion about new ESSA plans has included considerable examination of the use of an additional measure of student success or school performance that is explicitly contemplated by ESSA, but policymakers also have the opportunity to take more far-reaching and ambitious actions that would not only expand the measures used for outcome-based accountability but would also incorporate professional accountability systems.

### **Professional accountability merits more attention**

Professional accountability is prominent in many fields: professional organizations impose standards for entry, provide resources for continuing learning, and set standards of practice that may be monitored and enforced by direct observation of practice. Professional expectations in

medicine, for example make use of all four psychological mechanisms: doctors must pass a series of exams to be certified for practice (evaluation); board certifications for specializations are publicly reported (identifiability); medical rounds require doctors to explain cases and treatment plans to their colleagues (reason-giving); and surgery is conducted with other hospital staff present (mere presence of another). Most of these professional accountability tools not only maintain standards of practice but also help professionals learn how to improve their practice.

Professional accountability has long been used in K-12 education as well, but typically only to set unambitious minimum standards. Historically, standards of entry to the profession were low, collaboration among teachers was limited, and evaluations were pro forma—involving rare, brief observations of teachers by principals, who provided little useful feedback and a crude verdict that the teacher is "satisfactory" (or in very rare instances, not satisfactory) (Weisberg et al. 2009). The traditional "egg crate" organization of the school, in which individual teachers are alone in their classrooms, is not conducive to robust forms of professional accountability. One of the hallmarks of professional accountability is transparency of practice, which induces accountability through (at least) the mere presence of another—but teachers have typically been isolated from other adults, behind closed classroom doors.

More ambitious forms of professional accountability have appeared in schools in recent years. Evaluation of teachers and principals is taken more seriously, partly as a result of the efforts of major federal policy initiatives of the Obama/Duncan Administration (Troppe et al. 2017). States and districts have undertaken enormous efforts to improve and implement classroom observation instruments so that they provide better feedback for the improvement of professional practice (see, e.g., Lipscomb, Terziev, and Chaplin 2015; Chaplin et al. 2014). Professional learning communities that aim to increase collaboration among teachers are increasingly popular. In some

schools, teachers are observing each other and providing feedback without formal stakes attached (Furgeson et al. 2014). More principals are being evaluated based in part on feedback from their teachers (Wayne et al. 2016).

The rest of this paper explores a variety of professional accountability tools that might be used to promote continuous improvement in schools—focusing on tools that are already in use in innovative schools. The paper organizes the tools in three categories: transparency of practice, job redesign, and client feedback.

### **Professional accountability through transparency of practice**

Teacher evaluation systems have undergone major changes in most states in the country over the last several years (Troppe et al. 2017). Federal programs such as Race to the Top provided large incentives for states to revamp their educator evaluation systems so that they include valid, reliable, and comprehensive indicators of professional practice as well as measures of student achievement growth. These changes have become highly controversial, but the controversy has focused almost entirely on the use of student test scores in teacher evaluation: teachers, unions, and many parents have resisted the effort to push high-stakes testing to the level of individual teachers. The controversy over the use of test scores (and “value-added” statistical models that seek to identify a teacher’s contribution to student achievement growth) has obscured the fact that most of the effort in creating and implementing new evaluation systems has gone toward improving measures of teachers’ instructional practice—thereby making practice more transparent.

For observing classroom instruction, crude observation checklists have been replaced by detailed rubrics such as those produced by the Framework for Teaching, CLASS, and PLATO, which

expect evaluators to have substantial expertise in instructional techniques and require them to spend more time observing instruction than was typical in the pro forma observations of the past (Kane 2012; Troppe et al. 2017). These new rubrics seek to differentiate teachers across three to five rating categories (rather than merely “satisfactory” or not), to distinguish performance on multiple domains of instructional practice, and to provide useful feedback to help teachers improve their practice.

To be sure, there are concerns about the observation measures: it isn’t clear how much they distinguish instructional domains (Gill et al. 2016); rating scores tend to cluster near the high end of the scales (Lipscomb, Terziev, and Chaplin 2015); and they are costly and labor-intensive to implement (Dynarski 2016; Kraft and Gilmour 2016). Nonetheless, they are increasing the transparency of practice and creating accountability for practice that can be a useful counterweight to high-stakes tests. And there is evidence that they can produce improvements in student achievement (Steinberg and Sartain 2015).

Of course, in the context of formal evaluation, teacher observations involve high stakes too; they are distinguished from value-added measures not by the stakes but by the fact that they focus on what the teacher is doing rather than what the students have learned. But new observation rubrics can also be used—and can create accountability—outside the context of formal evaluation.

Increasing the rigor, intensity, and frequency of observation activates accountability through the mere presence of another. If the observations are accompanied by feedback and discussion, they also activate the reason-giving accountability mechanism. Both of these mechanisms create accountability even if the observations have no formal consequences.

Indeed, low-stakes transparency of practice may be most powerful if it is routine. In the Propel charter-school network in southwestern Pennsylvania, I have seen teachers not miss a beat when the principal enters the classroom, because they see the principal in classrooms every day to observe and support instruction, usually not in connection with any formal evaluation. Classroom doors are always open to visits from the principal, and teachers expect that the primary purpose of the visits is support; they recognize that the support requires transparency of practice.

Meanwhile, schools like Kansas City's Kauffman Charter School are *literally* increasing the transparency of instructional practice in a routine, low-stakes way by putting interior windows on their classrooms, so that any adult in the hallway can observe what is happening in class (Johnson et al. 2016). Schools that lack the resources to make major physical changes to their facilities could use video cameras to create similar transparency at lower cost. Video recordings are increasingly being used for formal observations of teaching, and video cameras might also be used in a more passive way that simply allows school principals and instructional support staff to have on-demand live access to each of the classrooms in a school. The creation of live video feeds from each classroom to the principal's office would create accountability that could not only support evaluation and feedback but also influence instruction through the mere presence of another. A more ambitious intervention might provide not only continuous live video access to relevant school staff, but also occasional snapshots of classroom video that could be made available to parents (after video processing that hides the faces of students, thereby maintaining necessary privacy protections).

Principal observations of teaching are necessarily limited by the number of hours in the day. Even if video recordings are used to allow the principal to "time shift" observations to after-school hours, principal observations are still constrained by the number of hours in the week and

by the principal's many competing responsibilities, especially in large schools. Moreover, any particular principal is likely to have incomplete knowledge of all of the content areas taught by different teachers, particularly at the secondary level. Relying exclusively on the principal to observe and provide feedback on teacher practice places superhuman expectations on the principal. Not surprisingly, research has found implementation in practice to fall short of the ideal (Kraft and Gilmour 2016).

But observations and feedback need not—and should not—be conducted only by principals. Some schools and districts are promoting ambitious forms of peer feedback by requiring regular observation by peers and/or participation of peers in the formal review process—with promising results for student outcomes. The Peer Assistance and Review system has been shown to increase the ability of reviewed teachers to promote student achievement growth (Taylor & Tyler 2012). Meanwhile, some schools are using peer observations not only for the benefit of the observed teacher but also for the observing teacher: New York City's TEP Charter School expects its teachers to observe each other twice every week, and it has produced large positive impacts on the achievement of its students (Furgeson et al. 2014). TEP believes that just as the observed teacher may benefit from feedback, the observer may learn effective teaching practices.

Instructional coaching by expert teachers has also gained adherents. Intensive instructional coaching is one of the distinguishing characteristics of a handful of charter-school management organizations that produce significant positive impacts on student achievement (Furgeson et al. 2012). Aspire, KIPP, Uncommon Schools, and YES Prep are all high-performing charter-school networks that rely on intensive coaching to promote strong and consistent instructional practice, particularly for novice teachers (Lake et al. 2012).

Ambitious forms of observation and feedback tend to be labor-intensive and expensive; some critics have doubted whether their benefits can justify the costs (Dynarski 2016). Even if substantial increases in observation by administrators and peers might be impossible or prohibitively expensive in some schools, more modest forms of professional accountability could create a degree of transparency at lower cost. For example, an intervention might activate the “reason-giving” accountability mechanism by expecting that teachers share lesson plans with each other once a week, provide quick comments on the lesson plans, and produce brief reflections on successful and unsuccessful elements of lessons.

To be sure, some teachers will not look favorably on accountability tools that make their practice transparent to administrators or to their colleagues. Indeed, one teacher’s public response to a suggestion that “Professional accountability means that teachers cannot close their doors and teach any way they please” (Gill and Lerner 2017) is an opposing view that “the teacher is the expert” who should “decide what is to be emphasized and how it is to be presented and taught” (comment responding to Gill and Lerner 2017). Although this kind of unconstrained individual autonomy is not how other professions actually work, it is not a unique view among teachers. Moreover, in some school districts, the provisions of union contracts may put obstacles in the way of peer observations. In one district where Mathematica assisted with developing a teacher evaluation system, longstanding practice had created a culture in which many teachers were unwilling to participate in evaluating their colleagues, even if the collective bargaining agreement could be modified to allow peer input.

It probably is not a coincidence that the examples of transparency of practice often come from charter schools. Charter schools are rarely subject to the constraints of collective bargaining agreements that might constrain the frequency of observation by administrators or preclude peer

involvement in observation and evaluation. Districts, in contrast, sometimes are burdened by longstanding mistrust and antagonism between teacher unions and central administration. In extreme cases, that can lead to provisions like the Chicago Public Schools' prohibition on using video recordings for teacher evaluations (Chicago Public Schools 2014)—an explicit rejection of transparency and the professional accountability that comes with it.

### **Professional accountability through job redesign**

Peer observations could be part of a larger professional accountability system that involves differentiating the responsibilities of effective teachers in ways that could extend their influence. In many countries that outperform the United States on international assessments, teachers are far more likely to be assigned mentors who provide feedback. Only 11 percent of American teachers reported receiving feedback from mentors on a recent international study of educational practice, compared to 24 percent in Australia, 38 percent in Singapore, and 39 percent in Japan—all of which outscored the United States in all three subjects (reading, math, and science in the latest study of the Program for International Student Assessment (OECD 2014a, 2014b).

Similarly, in three other high-performing countries, teachers were far more likely than American teachers to receive feedback from colleagues: 27 percent of American teachers, versus 84 percent of Korean teachers, 43 percent of Finnish teachers, and 57 percent of Dutch teachers. Moreover, teachers in these high-performing countries were more likely to report that the feedback they received led to public recognition, increased compensation, or career advancement (OECD 2014a).

Career advancement opportunities for teachers in American schools have typically been rare for those who do not wish to go into administration. The establishment of instructional coaching

positions is creating new career paths in some schools. This may create the opportunity for effective teachers to extend their influence by helping to develop the skills of their colleagues. But it also raises a question about tradeoffs: is it a good idea to take the best teachers out of their own classrooms to assist other teachers to improve their craft? At least one of the architects of an ambitious districtwide effort to create a new career ladder believes it was a mistake to put the most effective teachers in coaching positions (Franklin 2016). The appropriate balance is an empirical question that merits attention.

Schools that are implementing the Opportunity Culture program created by Public Impact are exploring other ways to promote professional accountability by differentiating teachers' jobs so that the most-effective teachers reach larger numbers of students (and are compensated more). Some schools participating in the Opportunity Culture initiative have organized teacher teams and changed student assignment policies so that the best teachers can serve as lead teachers for multiple classrooms, or take on larger class sizes, thereby allowing novice teachers to work with smaller numbers of students (Public Impact 2016). In principle, shifting class sizes so that better teachers serve more students could be an advantage for all: every student would have either a great teacher or a small class (Hansen 2014). Teaching more students would be a mark of professional expertise and would come with more salary. Budget simulations have shown that this could be done without the need for additional revenue (Roza and Warco 2015). The effectiveness of Opportunity Culture schools in promoting student achievement is currently being evaluated.

Teacher teams are also increasingly popular, particularly in the form of professional learning communities (PLCs). PLCs and other team-based initiatives seek to improve instruction by encouraging professional collaboration among teachers. TEP Charter School, for example, not

only expects teachers to observe each other regularly, but also has its teachers collaborate to create their own professional development for six weeks every summer. And it pays its teachers a base salary of \$125,000, which it affords in part by using larger class sizes than other New York City middle schools (Furgeson et al. 2014).

Giving teachers more (collective) responsibility in their schools is also contemplated by Ken Frank's "Constitution for Effective School Governance" (2012). Frank's proposal would give teachers explicit authority to approve or reject major changes to school programs. It would also give teachers some implicit leverage with the principal, who would be subject to possible negative career consequences if teachers exit the school in large numbers.

### **Professional accountability through input from clients**

Client voice is an implicit element of professional accountability in many fields. The rationale for client voice is even stronger in public education than in other endeavors, because school districts' traditional local monopoly on public education means that clients (parents and students) have only limited opportunities to select alternate providers. Options have increased in recent decades with the growth of charter schools in some communities, but choices remain limited for families who lack the resources to move to a different school district or pay private-school tuition. With limited exit options available, client voice is especially important.

Some educator evaluation systems are beginning to incorporate feedback from clients. Districts such as the Pittsburgh Public Schools are using student input as a component in teacher evaluations, just as colleges and universities do. To be sure, there is evidence from higher education that student ratings of teachers may be affected by implicit biases (MacNeill, Driscoll, and Hunt 2015). But they are considered to be critical input into instructor evaluation

nonetheless, and scores could be adjusted to account for implicit biases if necessary. In the K-12 context, survey results have been shown to be predictive of student achievement growth in the K-12 context (Kane 2012; Chaplin et al. 2014). And students' ability to assess the effectiveness of their teachers is not surprising, since, unlike principals or peers, students observe their teachers every single day. In Mathematica's work in school districts, we sometimes find teachers to be more suspicious of student input to evaluation even than they are of value-added measures of their contributions to achievement growth, perhaps due to concern that evaluations might become popularity contests. But well-designed surveys, such as the Tripod survey developed at Harvard University (Ferguson 2014), focus on students' experiences in the classroom rather than whether they like the teacher.

Client input is relevant for principals as well as teachers. Indeed, teachers are themselves arguably among the key clients (along with parents and students) of principals. Teacher input has not traditionally been a significant factor in principal evaluations, but it could be. Input from subordinates is commonly a part of "360" evaluations in the business world. The Vanderbilt Assessment of Leadership in Education (VAL-ED), now being used in hundreds of schools across the country, includes a survey of teachers to provide input in the principal's evaluation.

Like feedback from classroom observations, feedback from clients may have powerful accountability effects even if it is not tied to formal evaluation or otherwise has stakes attached. Even when the client input has no formal consequences, it makes use of the psychological accountability mechanism of identifiability, under the assumption that educators care about the perceptions of those they serve. The Institute of Education Sciences is currently conducting a randomized experimental evaluation of a no-stakes observation and feedback system that includes not only instructional feedback to teachers but also client feedback from teachers to

principals through the VAL-ED (Wayne et al. 2016). Similarly, student surveys could be used as an accountability and feedback tool for teachers even if they are not tied to formal teacher evaluations.

## **Conclusion**

It is past time for the term “accountability” in K-12 education policy to be liberated from its narrow association with high-stakes testing. As the behavioral science literature makes clear, accountability comes in many different forms. Policymakers’ single-minded focus on one specific form has led them to overlook opportunities to create a rich system of incentives and supports that would employ multiple accountability tools in concert. Professional accountability deserves particular attention, for two important reasons. First, it can activate all four psychological accountability mechanisms, including some (such as reason-giving and the mere presence of another) that are overlooked by outcome-based accountability (high-stakes testing). Second, it may be the only form of accountability that directly incorporates feedback to help educators improve their practice.

The examples described in this paper illustrate some of the wide variety of professional accountability tools that might be employed for school improvement. Initiatives that increase the transparency of educational practice, redefine instructional positions to extend the influence of the most effective teachers, and incorporate input from clients would enhance the accountability of educators and potentially improve school performance. By giving teaching a greater resemblance to other professions, such initiatives might also, over time, help to raise the status and esteem of the teaching profession.

**References**

- Chaplin, Duncan, Gill, B.P., Thompkins, A., & Miller, H. (2014). *Professional practice, student surveys, and value-added: Multiple measures of teacher effectiveness in the Pittsburgh Public Schools* (REL 2014–024). Washington, DC: U.S. Department of Education, Institute of Education Sciences, Regional Educational Laboratory Mid-Atlantic.
- Chicago Public Schools. (2014). *REACH Students educator evaluation handbook 2014–15*. Chicago: Author.
- Dynarski, Mark. (2016). “Teacher Observations Have Been a Waste of Time and Money.” Washington, DC: Brookings Institution.
- Ferguson, Ronald F. with Charlotte Danielson. (2014). “How Framework for Teaching and Tripod 7Cs Evidence Distinguish Key Components of Effective Teaching.” In Thomas J. Kane, Kerri A. Kerr and Robert C. Pianta, eds, *Designing Teacher Evaluation Systems: New Guidance from the Measures of Effective Teaching Project*. Hoboken, NJ: Jossey-Bass Publishers.
- Frank, Kenneth. (2012, February 24). Constitution for effective school governance [ID No. 16715]. *Teachers College Record*. Available from <http://www.tcrecord.org>
- Franklin, Samuel with Caitlin Blair. (2016.) “How to Improve Teaching: What Strategies Are Worth Pursuing.” Working paper, Carnegie Mellon University, Heinz College.
- Furgeson, Joshua, Brian Gill, Joshua Haimson, Alexandra Killewald, Moira McCullough, Ira Nichols-Barrer, Bing-ru Teh, and Natalya Verbitsky Savitz. (2012). “Charter-School Management Organizations: Diverse Strategies and Diverse Student Impacts.” Cambridge, MA: Mathematica Policy Research.
- Furgeson, Joshua, Moira McCullough, Clare Wolfendale, and Brian Gill. (2014). “The Equity Project Charter School: Impacts on Student Achievement.” Cambridge, MA: Mathematica Policy Research.
- Gill, Brian P., Jennifer Lerner, and Paul Meosky. (2016.) “Re-Imagining Accountability in K-12 Education.” *Behavioral Science and Policy* 2(1): 57-70.
- Gill, Brian P. and Jennifer Lerner. (2017). “Here’s What Works for Teacher Accountability.” *Education Week*, January 24. Available at [www.edweek.org](http://www.edweek.org).
- Gill, Brian P., Megan Shoji, Thomas Coen, and Kate Place. (2016). “The Content, Predictive Power, and Potential Bias in Five Widely Used Teacher Observation Instruments.” Washington, DC: U.S. Department of Education, Regional Educational Laboratory Mid-Atlantic.
- Hansen, Michael. (2014). “Right-sizing the Classroom: Making the Most of Great Teachers.” Washington, DC: CALDER Working Paper 110.

Johnson, Matthew, Cleo Jacobs Johnson, Scott Richman, Alicia Demers, Claudia Gentile, and Eric Lundquist (2016). "The Educational and Behavioral Impacts of the Ewing Marion Kauffman Charter School." Oakland, CA: Mathematica Policy Research Working Paper.

Kane, Thomas J. (2012). "Capturing the Dimensions of Effective Teaching." *Education Next* 12(4).

Kraft, Matthew A., David Blazar, and Dylan Hogan. (2016). "The Effect of Teacher Coaching on Instruction and Achievement." Working paper, Brown University.

Kraft, Matthew A. and Alison F. Gilmour. (2016). "Can Principals Promote Teacher Development as Evaluators? A Case Study of Principals' Views and Experiences." *Educational Administration Quarterly* 52(5): 711-753.

Kraft, Matthew A. and Alison F. Gilmour. (2017). "Revisiting the Widget Effect: Teacher Evaluation Reforms and the Distribution of Teacher Effectiveness." Working paper, Brown University.

Lake, Robin, Melissa Bowen, Allison Demeritt, Moira McCullough, Joshua Haimson, and Brian P. Gill. (2012). "Learning from Charter School Management Organizations: Strategies for Student Behavior and Teacher Coaching." Cambridge, MA and Seattle, WA: Mathematica Policy Research and the Center for Reinventing Public Education.

Lerner, Jennifer and Philip Tetlock. (1999). "Accounting for the Effects of Accountability." *Psychological Bulletin* 125: 255-275.

Lipscomb, Stephen, Jeffrey Terziev, and Duncan Chaplin. (2015). "Measuring Teachers' Effectiveness: A Report from Phase 3 of Pennsylvania's Pilot of the Framework for Teaching." Cambridge, MA: Mathematica Policy Research.

MacNell, Lillian, Adam Driscoll, and Andrea N. Hunt. (2015). "What's in a Name: Exposing Gender Bias in Student Ratings of Teaching." *Innovative Higher Education* 40(4): 291-303.

Organisation for Economic Co-operation and Development. (2014a). *TALIS 2013 results: An international perspective on teaching and learning*. Paris, France: Author.

Organisation for Economic Co-operation and Development. (2014b). *PISA 2012 results in focus*. Paris, France: Author.

Public Impact. (2016). "Pioneering Multi-Classroom Leaders: An Opportunity Culture Vignette Series." Available at [www.opportunityculture.org](http://www.opportunityculture.org).

Roza, Marguerite and Amanda Warco. (2015). "Paying the Best Teachers More to Teach More Students." Edunomics Lab at Georgetown University. Available at [www.edunomicslab.org](http://www.edunomicslab.org).

Steinberg, Matthew P. and Lauren Sartain. (2015). "Does Teacher Evaluation Improve School Performance? Experimental Evidence from Chicago's Excellence in Teaching Project." *Education Finance and Policy* 10(4): 535-572.

Taylor, Eric and John Tyler. (2012). "The Effect of Evaluation on Teacher Performance." *American Economic Review* 102(7): 3628-3651.

Troppe, Patricia, Anthony Milanowski, Camilla Heid, Brian P. Gill, and Christine Ross. (2017). "Implementation of Title I and Title II-A Program Initiatives: Results from 2013-14." Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Wayne, A. J., Garet, M. S., Brown, S., Rickles, J., Song, M., Manzeske, D. (2016). *Early Implementation Findings From a Study of Teacher and Principal Performance Measurement and Feedback: Year 1 Report* (NCEE 2017-4004). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. Washington, D.C.: The New Teacher Project.