

Reimagining Accountability in K-12 Education: A Behavioral Science Perspective

Faculty Research Working Paper Series

Brian P. Gill

Mathematica Policy Research

Jennifer S. Lerner

Harvard Kennedy School

Paul Meosky

Harvard University

April 2016

RWP16-018

Visit the **HKS Faculty Research Working Paper Series**

at: <https://research.hks.harvard.edu/publications/workingpapers/Index.aspx>

The views expressed in the **HKS Faculty Research Working Paper Series** are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). Papers may be downloaded for personal use only.

Reimagining Accountability in K-12 Education: A Behavioral Science Perspective

Brian P. Gill
Mathematica Policy Research

Jennifer S. Lerner
Harvard University

Paul Meosky
Harvard University

DRAFT, 2016.02.15

Accepted for publication in *Behavioral Science & Policy*

Contact: Brian P. Gill, bgill@mathematica-mpr.com

The authors thank Phil Tetlock, Mark Dynarski, Sandy Jencks, Jenny Mansbridge, Mark Moore, Todd Rogers, Christina LiCalsi, Katie Shonk, Paul Hill, Ken Frank, and the participants in the seminar series of the University of Arkansas' Department of Education Reform for helpful comments on earlier versions of this paper.

Abstract

The primary lever American policymakers have used to improve school performance is “accountability” in the form of high-stakes testing. But the behavioral literature, overlooked in the education policy debate, shows that accountability exists in a variety of forms that evoke different psychological mechanisms and can have positive or negative effects. Examining the psychological/behavioral literature alongside the education literature, we identify four forms of accountability relevant to K-12 schooling: outcome-based (high-stakes testing), rule-based, market-based, and professional accountability. Promoting continuous improvement in schools is likely to require multiple forms of accountability that not only offer rewards and sanctions but also increase the transparency of educational practice and provide mechanisms for improving practice. This suggests that professional accountability—which has historically been underutilized in schools—merits particular attention.

The No Child Left Behind Act of 2002 (NCLB) required reporting of every public school's level of student proficiency in math and reading, and imposed sanctions for failure to achieve targets. Schools that fell short of state-determined proficiency goals faced mandatory interventions; chronically failing schools could be re-staffed, taken over by the state, or shut down. In education policy, "accountability" came to be synonymous with high-stakes testing, which became the primary policy lever to improve school performance. Under the leadership of Arne Duncan, the Obama Administration's Department of Education doubled down on high-stakes testing, aggressively encouraging states to include student achievement growth in the evaluation of teachers and principals.

High-stakes testing is now under attack on multiple fronts. Unions have objected to the use of test scores in teacher evaluations. Scholars have proposed refining or replacing the current accountability regime.^{1 2 3} An "opt out" movement has enlisted parents who refuse to have their children take the standardized tests. In December, recognizing the increasing unpopularity of the federal mandates for high-stakes testing, Congress passed the Every Student Succeeds Act, which replaces NCLB and gives states far more discretion to design their own accountability policies.

With much more latitude about the design of accountability policies, policymakers need to be informed about the empirical evidence on accountability's effects. Unfortunately, the education policy debate has defined accountability to have a particular, narrow meaning that could lead policymakers to overlook important tools for improving school performance. An extensive literature in experimental social psychology identifies the conditions under which specific forms of accountability can improve outcomes, have no effect, or make matters worse. This literature, which could provide valuable insights for the design of school accountability regimes, has been largely overlooked in the education policy debate.⁴

Lerner and Tetlock's comprehensive review of the psychological literature on accountability⁵ makes two broad points that motivate our discussion. First, accountability comes in many forms, involving different mechanisms and different behavioral responses. Outcome-based accountability (such as high-stakes testing) is only one of these forms. Second, accountability can have positive and/or negative effects on judgments and decision making, depending on the accountability type, the decision context, and the nature of the task. In the context of schooling, this means that policymakers have more options than just high-stakes testing and the absence of accountability.

We examine the behavioral literature alongside the education literature to inform a consideration of accountability regimes in education, discussing evidence related to four types of accountability: rule-based, market-based, outcome-based, and professional accountability. and propose how policymakers might use a range of accountability tools to promote continuous improvement in schools. Some of the other forms of accountability are already in use in limited ways in schooling, but they are typically not identified as part of an accountability system. Complicating the classic principal-agent problem that accountability policies seek to address, schooling has multiple public and private purposes with multiple constituencies, including parents, students and the public. Any single form of accountability on its own is unlikely to fully serve these multiple purposes and constituencies.

We reach several conclusions relevant to policy and practice. First, transparency alone can create accountability even in the absence of explicit rewards and sanctions. Second, multiple forms of accountability can be used in conjunction to take advantage of complementary strengths and weaknesses. Third, to promote continuous improvement in schools, a comprehensive

accountability system should include mechanisms for the improvement of practice. Collectively, these three points suggest greater use of professional accountability than has historically been the norm in K-12 education, alongside rule-based, market-based, and outcome-based accountability.

I. Defining Accountability

Lerner and Tetlock's review identifies four mechanisms that make people feel accountable: (1) the *mere presence of another*—simple awareness that someone else is watching; (2) *identifiability*—the expectation that an action or outcome will be personally attributable; (3) *reason-giving*—the expectation that one will need to explain actions; and (4) *evaluation*—the assessment of one's performance with particular rules and consequences.

The outcome-based (high-stakes testing) accountability regime inaugurated by NCLB incorporates identifiability and evaluation, but does not recognize that there are also less-aggressive ways to create accountability, through reason-giving or the mere presence of another. In particular, policies and practices that increase transparency, making the relevant activity more visible to others, may evoke any or all of the first three accountability mechanisms even without imposing formal consequences. Other professions—from law to engineering to architecture to medicine—typically use multiple forms of accountability that collectively make use of all four of the accountability mechanisms identified by Lerner and Tetlock.

One common form is *rule-based accountability*, which requires decision makers to act in accordance with rules that delineate mandated or forbidden activities. Rule-based accountability relies on the identifiability of actors and sometimes includes an evaluation component. Rule-based accountability is common in education as in other fields: state education codes, regulations, and teacher contracts create rule-based accountability.

Market-based accountability is pervasive in most fields; lawyers and engineers are accountable to their customers in a way that public-school educators typically are not. Under market-based accountability, clients or customers are the primary constituency, and they implicitly hold decision makers responsible by choosing among providers. Market-based accountability employs the identifiability mechanism, and it encourages providers to describe and explain their services, thereby invoking reason-giving as well. Market-based accountability is relevant in private schools, public charter schools, and in a limited way in conventional public schools (to the extent that families have the ability to move to a desirable school district).

In many fields, *professional accountability* systems go beyond what markets and government regulations require. Professional organizations impose standards for entry, provide resources for continuing learning, and set standards of practice that may be enforced by direct observation of practice (such as when medical residents are observed by senior physicians). These approaches can evoke all four of the accountability mechanisms identified by Lerner and Tetlock. K-12 education includes some forms of professional accountability, for example in certification requirements, but the professional accountability in K-12 education typically has not involved much more than modest requirements for entering the profession and minimal requirements for maintaining professional status.

The kind of *outcome-based accountability* that has been the focus of education policy over the last two decades has not generally been common in other fields, perhaps because market-based accountability serves the same function. Tort law, in which plaintiffs' attorneys are paid only if they win, is a notable exception. Outcome-based accountability is increasingly being

attempted in health care⁶, which resembles K-12 schooling in that consumers have limited market power.

In sum, various forms of accountability operate in various fields, and the different forms of accountability evoke different psychological mechanisms. Table 1, below provides an overview of the different mechanisms evoked by different accountability types, with examples (outside of K-12 education) in each relevant cell.

Table 1. Accountability types in policy and psychological accountability mechanisms

		Accountability types in policy			
		Outcome-based	Rule-based	Market-based	Professional
Psychological accountability mechanisms	Mere presence of another				Surgical operating room with nurses in attendance
	Identifiability	US News college rankings	Minimum certification requirements (various professions)	Consumer Reports, Better Business Bureau	Membership in professional organization
	Reason-giving			Annual report to company stockholders	Medical rounds with explanation of treatment
	Evaluation	Contingent fees for attorneys	Driver licensing test		Bar exam

Clearly, policymakers have more tools than just outcome-based accountability to promote school performance. Indeed, outcome-based accountability systems ignore two of the four psychological accountability mechanisms identified by Lerner and Tetlock, thereby leaving tools for improvement unused. If the “presence of another” and “reason-giving” mechanisms are to be used for school improvement, they will require accountability approaches other than high-stakes testing. The rest of this paper considers applications of all of these accountability types, using research from the field and the laboratory to inform ways that K-12 accountability regimes might be designed to improve educational outcomes.

II. Outcome-based Accountability

Twenty-five years ago, outcome-based accountability was almost unknown in K-12 schooling. The education standards movement that gained steam during the 1990s promoted clear expectations at each grade level and tests to measure students’ “proficiency.” Beginning in 1994, federal law required states to set proficiency standards, assess their students in multiple grades, and report school-specific results. Thus, the primary accountability mechanism was identifiability, in the public reporting of results. Eight years later, NCLB added explicit sanctions to schools falling short of proficiency targets, thereby supplementing identifiability with evaluation. More recently, the federal government has pushed states to extend outcome-based accountability from schools to individual educators by including test-score growth in the evaluations of teachers and principals.

Behavioral Evidence on Outcome-based Accountability. In randomized experiments, the behavioral literature has found few instances of positive effects of outcome-based accountability. Because it does not constrain decision makers with rules, outcome-based accountability might be more effective than other forms of accountability at promoting innovation⁷, but this has not been extensively examined empirically. The effectiveness of outcome-based accountability can be undermined by the sunk-cost bias, which makes decision makers more likely to pursue action because of prior investments, even with low odds of success.⁸ In addition, tangible rewards sometimes undermine intrinsic motivation.⁹

By increasing a decision's difficulty, outcome-based accountability may impair decision making by eliciting stress and negative emotions.^{10 11} Perhaps because stress burdens cognition^{12 13}, outcome accountability sometimes fails to increase the use of strategies that require substantial effort.⁵

In addition, accountability regimes viewed as illegitimate can be counterproductive.⁵ Many teachers are suspicious of "value-added" models (VAMs) that quantify contributions to student achievement. Improving on cruder outcome-based accountability regimes that rely on student achievement *levels*, VAMs account statistically for factors outside the teacher's control. Although well-designed VAMs can produce unbiased measures of teachers' contributions to student achievement growth^{14 15}, suspicions on the part of practitioners could undermine their ability to promote performance improvements.

Moreover, even unbiased measures of educators' contributions to student achievement can be problematic in a high-stakes accountability regime. Tests cannot capture all of the skills and knowledge that schools seek to impart. Some evidence suggests that instructional practices that raise test scores differ from those that promote students' effort and long-term goals.¹⁶ High stakes encourage schools to focus on tested elements of learning to the detriment of untested elements, as demonstrated by the psychology literature on conformity and the availability heuristic¹⁷; and may undermine the validity of the test itself.¹⁸

Field Evidence on Outcome-based Accountability. The effects of NCLB in particular and high-stakes testing in general are hotly debated. Most existing studies suggest positive effects in at least some schools, grades, and subjects.^{19 20 21 22} The effects of performance-pay incentives for teachers on student achievement have been mixed.^{23 24 25}

Meanwhile, many schools have narrowed the curriculum²⁶ and spent increasing time preparing for state assessments.^{27 28} Scores on low-stakes assessments have improved more slowly than scores on high-stakes assessments.^{29 30} In extreme cases, educators have been caught cheating. Teacher-developed "student learning objectives" (SLOs), increasingly used as outcome-based accountability measures, may be especially susceptible to inflation, because they ask teachers to grade themselves.³¹

III. Rule-based Accountability

Rule-based accountability regimes implicitly rely on identifiability and sometimes evaluation. Historically, schools have relied on rule-based accountability to define structural conditions of schooling and set minimum expectations.³² Rule-based accountability has included state-level decisions about textbooks; contractual rules about working conditions, hours, and class size; federal and state spending regulations; and traditional, pro forma teacher evaluations, which typically deem 98-99% of teachers satisfactory.³³ Teachers typically had wide discretion

about instruction throughout the 20th century³⁴, which permitted wide variation in effectiveness. Rules and protocols may have ensured minimum standards and reduced the most egregious inequities, but they have reduced opportunities for innovation and may have sacrificed efficiency for regularity.

Recently, some districts and school management organizations have become more directive about elements of instruction and school operations, pursuing a maximal version of rule-based accountability in which all teachers of the same courses may be expected to cover the same material at the same pace.³⁵ Principals have been asked to take on greater responsibilities as instructional leaders. Pacing guides are commonly used, and instruction is tied to state standards, with some lesson plans scripted to the minute.³⁶

Limited evidence supports a maximal rule-based approach in schools. Scripted “direct instruction” has been found to promote student achievement in elementary grades.³⁷ But maximal rule-based accountability could become counterproductive, because close monitoring often reduces intrinsic motivation.³⁸ Intense procedural scrutiny can also exacerbate the sunk-cost bias³⁹, undermine innovation, and entrench suboptimal practices.⁴⁰ The perception of rules as illegitimate can produce a boomerang effect.⁴¹ Indeed, aggressive rule-based accountability may be especially unsuited to teaching, an inherently complex task that requires daily adjustments and judgments.

IV. Market-based Accountability

Market-based accountability is based on classical economic principles rather than newer behavioral approaches. It involves the identifiability and reason-giving mechanisms: schools chosen by families must be identifiable and attractive to parents. Historically, market-based accountability did not play a substantial role in U.S. public education. Operating alongside tuition-based private schools, the public system has been based on the “common school” model, which assumes that each community will educate its children together, with school districts maintaining local monopolies on publicly supported education.^{42 43}

Policymakers have shown increasing interest in incorporating market-based accountability into education, reasoning that local monopolies controlled by school boards may not produce the best schools^{44 45}, and that giving families choices in schooling is inherently valuable.⁴⁶ These views have led to the rise of charter schools—publicly funded schools of choice that operate outside the direct control of local school districts.⁴⁷ The first charter schools opened only two decades ago; today, over 6,000 operate in more than 40 states. Meanwhile, a smaller number of states have adopted voucher programs that permit students to attend private schools at public expense.

Empirically, the evidence on the effects of market-based schools on students’ test scores and longer-term educational attainment suggests promise but is not definitive. In some contexts and locations, charter schools are producing substantial positive effects^{48 49 50 51}, but their performance varies widely.^{52 53} A few studies of small-scale voucher programs have found positive educational impacts, particularly for African-American students.^{54 55}

K-12 schooling differs from other services in ways that might make exclusive reliance on the market suboptimal. First, the classic principal-agent problem—aligning the interests of agents (educators) and clients—is complicated by the involvement of multiple clients (students, parents, and the public), whose interests may not be fully aligned. Moreover, children are presumed to be not fully capable of knowing their own interests. In addition, students’

educational experiences are affected not only by school quality, but by externalities, including characteristics of other students.⁵⁶ As a consequence, an unfettered market may produce segregated schools, as parents with high levels of knowledge, wealth, or motivation seek out schools with other families like their own.⁵⁷

Relatedly, skeptics worry that market-based schools will drain conventional public schools of funding and motivated families. Supporters argue that breaking the local monopoly produces healthy competitive pressure that will benefit all students. Although several studies have found neither positive or negative effects of charter schools on achievement in nearby conventional schools^{57 58 59}, others have found evidence of small positive effects^{59 60 61}; only one has found a negative effect.⁶²

Another externality relates to the original rationale for public education: society benefits from the inculcation of the knowledge, skills, and attitudes necessary for effective citizenship in a democracy. Historically, this key rationale for the common-school model implicitly assumed that effective education for citizenship required public operation of schools.^{42 43} The fact that the education of citizens is a public good argues against relying entirely on market mechanisms.

In fact, existing school choice programs rarely rely exclusively on market accountability. Charter schools are exempt from some forms of regulation but, like conventional public schools, are subject to rule-based and outcome-based accountability: They cannot charge tuition, they must submit their students to the same high-stakes tests taken in conventional public schools, and they (typically) must admit all applicants for whom they have space. Moreover, charter schools operate under the supervision of publicly empowered authorizers. Even the private schools that participate in voucher programs typically must submit to some regulation to receive public funds. Milwaukee's program, the longest-operating publicly funded voucher program in the country, imposes requirements for instructional time, forbids tuition, requires state assessments, and does not allow selective admissions.

V. Professional Accountability

Prominent voices are calling for professional accountability that would give teachers support, collaboration, and training, and set higher expectations.^{1 2} Professional accountability can take many forms, involving all four of the accountability mechanisms identified by Lerner and Tetlock. Licensing, standards, and professional reviews involve evaluation; observations and assistance by supervisors, instructional coaches, peers, or mentors involve identifiability, reason-giving, and sometimes evaluation; collaboration and co-teaching involve the presence of another and reason-giving.

Professional accountability is not synonymous with professionalism. Professionalism implies an ethic of meeting standards even in the absence of observers and consequences. Thus, we do not consider professionalism per se to be a form of accountability, which by definition involves an external observer.

Traditional and Novel Versions of Professional Accountability. States have long applied requirements for teaching licenses, including coursework, student teaching, and passing exams. Teacher contracts generally reward master's degrees and experience as proxies for professional skill. But master's degrees have little or no relationship to improved student achievement^{63 64}, and most studies find no effect of professional development on student achievement.^{65 66}

More robust and ambitious forms of professional accountability may hold more promise. Licensing and professional requirements at a high enough level—such as the certification process of the National Board for Professional Teaching Standards—might help identify especially effective teachers.⁶⁷ Teacher quality is the most influential school-controlled factor in student achievement growth⁶⁸, a fact that has prompted policymakers' aggressive promotion of more rigorous teacher evaluations. States and school districts are adopting extensive rubrics for the observation and rating of teaching practice. The new systems could deteriorate into compliance exercises that resemble traditional rule-based accountability mechanisms. But if they function as intended, they could substantially improve practice. Robust professional accountability systems—unlike outcome-based, rule-based, and market-based accountability—include tools and resources to help teachers improve. If taken beyond screening and compensation reform, they have the unique advantage of coupling accountability with support.

Novel forms of professional accountability might include new job descriptions and training methods. Some school districts have recently created teacher residency programs modeled on medical residencies, in which aspiring teachers spend much more time in the classroom. Other districts and schools are giving teachers leadership opportunities to develop the skills of their colleagues⁶⁹, as through instructional coaching.⁷⁰

In most professions, professional accountability includes answerability to clients.⁷¹ Unlike in higher education, K-12 schooling traditionally involves little direct accountability of educators to students. School districts such as the Pittsburgh Public Schools are now including student surveys in new teacher evaluation systems. And some are using teacher surveys in principal evaluation⁷², applying the business world's 360-degree feedback. Frank's "Constitution for Effective School Governance" would give every school a governing board of community members, to which the principal would be answerable.⁷³

Another professional accountability system is an intensive review of school quality conducted by independent, expert educators, as is common in British schools. A school quality review involves an extended visit by outside experts who observe instruction; interview teachers, students, and parents; and examine school performance data.^{1, 2} The review concludes with clear recommendations for improvement.

Professional accountability would make teaching more transparent, potentially activating all four accountability mechanisms. Indeed, rich professional accountability systems emphatically reject the notion that professionalism means allowing teachers complete discretion to practice as they choose. Rules such as those found in the Chicago Public Schools⁷⁴ that explicitly prohibit the use of classroom video recordings for evaluation are anathema to this vision of professional accountability. In contrast, schools like the Kauffman Charter School in Kansas City are literally making teaching transparent by giving classrooms interior windows that make them visible to adults in the hallways.

Greater teaching transparency is common in some countries that consistently outperform the United States in international comparisons of student achievement. A recent international study of educational practice found that although responding American teachers were more likely than average to receive feedback from principals, only 11% received feedback from mentors, versus 39% in Japan, 38% in Singapore, and 24% in Australia.⁷⁵ And only 27% of responding American teachers received feedback from colleagues, versus 84% in Korea, 57% in the Netherlands, and 43% in Finland. Those countries significantly outperformed the U.S. in math, reading, and science in the most recent study of the Program for International Assessment. American teachers were also far less likely than their counterparts in other countries to receive

feedback from student surveys, and less likely to report that the feedback they received led to public recognition, career advancement, or increased compensation.

Indeed, one study in an American urban school district found that improvements in student achievement were associated with teams of teachers that had stronger mutual professional ties, and with individual teachers who had stronger ties with their principals.⁷⁶ Professional accountability would promote ties among teachers, potentially developing the social capital and trust that have been found to be markers of effective schools.^{76 77}

Behavioral Evidence on Professional Accountability. Many studies demonstrate favorable effects of requiring people to justify their decisions to others, a common expectation of professional accountability regimes. One study found that requiring subjects to justify their decisions encourages high-effort strategies that are sensitive to evidence, alleviating mistakes and inconsistencies.⁷⁸ Similarly, another study found that the need to justify decisions stimulated systematic thinking and attention to evidence.⁷⁹ In a third study, a justification requirement reduced reliance on stereotypic thinking.⁸⁰ Subjects who had to justify their judgments have also been found to be less likely to be led astray by the fundamental attribution error (the tendency to over-attribute responsibility to individuals rather than situations).⁸¹ And process accountability for groups has been found to increase demand for information, induce information sharing, and produce better outcomes.⁸²

Other studies suggest that an increased sense of control (which might be promoted via a professional accountability regime that promotes initiative) may improve performance on attention-demanding tasks, promote more considerate decision-making, and assist memory formation.^{83 84} More generally, professional accountability may best encourage the systematic, effortful, and self-critical thinking associated with even-handed, accurate reasoning using systematic (rather than heuristic) processing.⁸⁵ Professional accountability might offer the flexibility needed for innovation while disallowing the adoption of the cognitively lazy but easily defensible decision.⁸⁶

Professional accountability is also compatible with the behavioral “nudges” that are increasingly being adopted in various areas of public policy.⁸⁷ Field trials have demonstrated, for example, that appealing to social norms (“most people like you do X,” following the work of Robert Cialdini⁸⁸) powerfully influences behavior in contexts ranging from tax collection⁸⁹ to motivating parents to keep their children in school.⁹⁰ This suggests the possibility that evaluation feedback to teachers, for example, could lead to improvement even in the absence of explicit consequences, if relevant information is provided appropriately.

Even though professional accountability is compatible with low-cost behavioral nudges, many forms of professional accountability are expensive and/or make substantial demands on educators. Teachers need time to observe each other and provide feedback. Instructional coaches need to be hired. School quality reviews must be staffed. More research will be needed to assess whether some forms of professional accountability are more cost-effective than others.

Field Evidence on Professional Accountability. A few studies have examined new forms of professional accountability for educators. Teacher residencies are showing promise in producing high-performing teachers and keeping them in the classroom⁹¹, and early evidence on the effects of instructional coaching on student achievement is encouraging.^{92 93 94} Several recently developed rubrics for observing and evaluating instructional practice have produced

evaluation ratings that are correlated with teachers' value added^{95 96 97}, suggesting that careful observation can produce feedback that could improve student outcomes.

In higher education, student evaluations of teachers have had positive effects on teaching.⁹⁸ Recent studies examining student surveys in K-12 schools have found the results are (modestly) correlated with measures of teachers' contributions to student achievement^{95 96}, suggesting promise.

An intensive, peer-based teacher evaluation system used in Cincinnati offers encouraging evidence on formal, job-embedded professional accountability. Participating teachers substantially increased their effectiveness in raising student achievement during and after the year they were evaluated by peers—even though the evaluation criteria were based entirely on professional practice, not on test results.⁹⁹ Dee and Wyckoff found evidence of favorable effects for low-rated and high-rated teachers on the teacher evaluation system now used by Washington, D.C. public schools.¹⁰⁰ And one experimental study found that “nudging” school principals with information about teachers' effectiveness increased the exit rate of low-value-added teachers and raised test scores, even without incorporating the information in a formal, high-stakes evaluation measure.¹⁰¹

VI. Conclusion: Increasing Professional Accountability and Transparency in a Multi-mode Approach

The term “accountability,” as it is used in K-12 education policy, reflects an unnecessarily narrow understanding of the range of accountability mechanisms that are available to policymakers and that can promote desired practices in schools. The outcome-based accountability that has been the focus of policymakers' attention can produce positive results, but relying on it excessively without a balance of other forms of accountability is likely to produce unintended and undesirable consequences.

Rule-based accountability, long included in American education policy, has helped set minimum standards and expectations, but has not been designed to promote high performance. And the effectiveness of more aggressive rule-based approaches is likely to decline with the decision maker's increasing distance from the classroom. Detailed mandates from distant officials are especially susceptible to being perceived by educators as illegitimate. Even though rule-based accountability has an important role in setting minimum standards, it is likely to have limited value in promoting continuous improvement in educational performance.

Market-based accountability likewise can play a productive role in improving school performance, but it is unlikely to be sufficient on its own. Given the imperfect information available to parents, the spillover effects of student sorting, and the public purposes of schooling, rule-based constraints on market-based accountability make good sense.

Professional accountability may help to fill the gaps. It is the most underused form of accountability in K-12 schooling in America. Barriers to entering the profession are low. Evaluation standards have historically been low. Compensation and career advancement have not been based on performance. And there has been little expectation that teachers will continually improve their practice, or even that their practice will be regularly observed by peers. This is changing, with initiatives related to rigorous educator evaluation and job-embedded professional development (coaching) now underway in schools across the country.

Below is a modified version of the table the beginning of the paper. As in the original table, we provide examples illustrating how different accountability policy types can evoke

different psychological accountability mechanisms. In the revised table (Table 2), however, the examples are specifically relevant to K-12 education, confirming that policymakers and educators have a wide range of accountability tools to employ.

Table 2. 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 ratings by local news media	Peer observation
	Reason-giving			Charter-school enrollment fairs	Coaching
	Evaluation	Value-added incorporated in teacher evaluations	Formal observation by principal		Advanced certification

After reviewing the evidence from behavioral science and the field literature, we propose that policymakers designing K-12 accountability systems should consider three key points. *First*, transparency alone can create accountability, even in the absence of explicit rewards and sanctions. Transparency of practice can activate several psychological accountability mechanisms (presence of another, identifiability, and reason-giving) that can powerfully influence behavior. Transparency also creates opportunities to inform the improvement of practice through feedback. And transparency can be promoted by professional accountability in various ways, including peer observation and evaluation, instructional coaching, and “360”-type feedback for principals and teachers alike.

Second, the diversity of advantages and disadvantages of different forms of accountability suggests that *multiple forms of accountability* might be usefully employed in complementary ways. Patil, Vieider, and Tetlock recently noted that outcome-based accountability may better promote innovation, while process-based accountability (including forms of professional accountability) may better promote the use of identified best practices.⁴⁰ They suggest that the disadvantages of both types might be counteracted by systems that promote the empowerment of decision makers to reduce conformity to deficient standard practices, encourage focus on outcomes, and facilitate organizational learning.¹⁰² Empowerment is implicit in professional accountability but can also be incorporated in an outcome-based accountability regime that communicates a desire to achieve shared objectives, they argue. Similarly, according to a National Academies report on high-stakes testing, external rewards are most likely to be effective when they are well aligned with educators’ intrinsic aims, promoting “autonomous motivation.”⁴ Creative use of behavioral nudges permits professional discretion while encouraging desirable practices.

Frank has proposed a form of school governance that would employ multiple modes of accountability, making the principal accountable to a community board and giving the principal

greater authority to remove the lowest-performing teachers, while at the same time giving teachers more say in school operations (including evaluation of the principal).⁷³ The proposal aims to improve teaching effectiveness by promoting collaboration among teachers.

The Equity Project (TEP) Charter School in New York City provides a compelling example of the advantages of using multiple accountability approaches. The school pays its teachers \$125,000 plus bonuses based on school-wide improvements in student achievement, but its approach goes beyond compensation. TEP's teachers are rigorously screened; have contracts renewed based on performance; engage in six weeks of professional development annually; and observe each other in the classroom at least twice a week, providing written feedback to their colleagues. In short, TEP's model includes professional accountability alongside market accountability and substantial outcome accountability. Moreover, TEP includes all of these accountability approaches while spending only the standard per-pupil funding allocation given to any New York City charter school. It has found a way to support both high pay and robust professional accountability by reallocating resources rather than raising costs. In its first years of operation, TEP has produced substantial positive effects on student achievement.¹⁰³

Third, employing multiple forms of accountability and multiple measures of practice and outcomes helps to create a complete system that provides *mechanisms to promote the improvement of practice*. As a recent study of accountability in public-sector organizations found, organizational learning requires feedback for improvement.¹⁰² Transparency of practice creates opportunities for educators to improve; rich data on student outcomes help diagnose students' needs; and rewards for success allow educators to innovate in productive ways.

In sum, improving educational effectiveness will likely require multiple accountability approaches in a package that creates transparency and empowers educators in the service of achieving positive student outcomes. And that will require shifting the balance toward a larger role for professional accountability.

Sources cited

¹ Tucker, M. S. (2014). Fixing our national accountability system. Washington, DC: National Center on Education and the Economy.

² Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives*, 22 (86).

³ Center on Reinventing Public Education. (2014). Designing the next generation of state education accountability systems. Retrieved from www.crpe.org.

⁴ A National Academies report examined some of the behavioral literature related to the use of test-based incentives, but did not examine other forms of accountability. Hout, M., & Elliott, S. W. (Eds.). (2011). *Incentives and test-based accountability in education*. Washington, DC: National Academies Press.

⁵ Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125, 255–275.

⁶ Gold, M. (2010). *Accountable care organizations: Will they deliver?* Princeton, NJ: Mathematica Policy Research.

⁷ Patil, S. V., & Tetlock, P. E. (2014). Punctuated incongruity: A new approach to managing trade-offs between conformity and deviation. *Research in Organizational Behavior*, 34, 155–171.

⁸ Simonson, I., & Staw, B.M. (1992). De-escalation strategies: A comparison of techniques for reducing commitment to losing courses of action. *Journal of Applied Psychology*, 77(4), 419–426

-
- ⁹ Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin*, *125*(6), 627–688.
- ¹⁰ Zhang, Y., & Mittal, V. (2005). Decision difficulty: Effects of procedural and outcome accountability. *Journal of Consumer Research*, *32*, 465–472.
- ¹¹ Siegel-Jacobs, K., & Yates, J. F. (1996). Effects of procedural and outcome accountability on judgment quality. *Organizational Behavior and Human Decision Processes*, *1*, 1–17.
- ¹² Mendl, M. (1999). Performing under pressure: Stress and cognitive function. *Applied Animal Behaviour Science*, *65*(3), 221–244.
- ¹³ Lazarus, R.S., & Folkman, S. (1984). *Stress, appraisal, and coping*. New York: Springer.
- ¹⁴ Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). *Have we identified effective teachers? Validating measures of effective teaching using random assignment*. Seattle, WA: Bill & Melinda Gates Foundation.
- ¹⁵ Chetty, R., Friedman, J.N., & Rockoff, J.E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9): 2593-2632.
- ¹⁶ Ferguson, R. F., with Danielson, C. (2014). How Framework for Teaching and Tripod 7Cs evidence distinguish key components of effective teaching. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the Measures of Effective Teaching project*. Hoboken, NJ: Jossey-Bass Publishers.
- ¹⁷ Tetlock, P. E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology*, *57*(4), 632–640.
- ¹⁸ Campbell, D. T. (1976). *Assessing the impact of planned social change* (Occasional paper no. 8). Hanover, NH: Dartmouth College Public Affairs Center.
- ¹⁹ Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, *30*(3), 418–446.
- ²⁰ Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, *24*(4): 305–331.
- ²¹ Ahn, T., & Vigdor, J. (2014). The impact of No Child Left Behind’s accountability sanctions on school performance: Regression discontinuity evidence from North Carolina (Working Paper no. 20511). Cambridge, MA: National Bureau of Economic Research.
- ²² Deming et al. find mixed results in different schools. Deming, D. J., Cohodes, S., Jennings, J., & Jencks, C. (2013). School accountability, postsecondary attainment, and earnings (Working paper no. 19444). Cambridge, MA: National Bureau of Economic Research.
- ²³ Chiang, H., Wellington, A., Hallgren, K., Speroni, C., Herrmann, M., Glazerman, S., & Constantine, J. (2015). *Evaluation of the Teacher Incentive Fund: Implementation and Impacts of Pay-for-Performance After Two Years*. Washington: U.S. Department of Education, Institute of Education Sciences.
- ²⁴ Springer, M. G., Pane, J. F., Le, V., McCaffrey, D. F., Burns, S. F., Hamilton, L. S., & Stecher, B. (2012). Team pay for performance: experimental evidence from the Round Rock Pilot Project on Team Incentives. *Educational Evaluation and Policy Analysis*, *34*(4), 367–390.
- ²⁵ Glazerman, S., & Seifullah, A. (2012). *An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) After Four Years*. Washington, DC: Mathematica Policy Research
- ²⁶ Dee, T. S., Jacob, B., & Schwartz, N. L. (2013). The effect of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis*, *35*(2), 252–279.
- ²⁷ Hamilton, L. S., Stecher, B. M., & Yuan, K. (2012). Standards-based accountability in the United States: Lessons learned and future directions. *Education Inquiry* *3*(2), 149–170.
- ²⁸ Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, *87*(2), 125–141.
- ²⁹ Holcombe, R., Jennings, J., & Koretz, D. (2013). The roots of score inflation: An examination of opportunities in two states’ tests. In G. Sunderman (Ed.), *Charting reform, achieving equity in a diverse nation* (pp. 163–189). Greenwich, CT: Information Age Publishing.
- ³⁰ Jacob, B. (2005). Accountability, incentives, and behavior: The impact of high-stakes testing in the Chicago public schools. *Journal of Public Economics*, *89*, 761–796.
- ³¹ Gill, B., English, B., Ferguson, J., & McCullough, M. (2014). *Alternative student growth measures for teacher evaluation: profiles of early-adopting districts*. Washington, DC: U.S. Department of Education, Institute of
-

Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic.

³² Tyack, D. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.

³³ Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness*. New York, NY: The New Teacher Project.

³⁴ Tyack, D. & Cuban, L. (1995). *Tinkering Toward Utopia: A Century of Public School Reform*. Cambridge, MA: Harvard University Press.

³⁵ Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban school systems improve student achievement*. New York: MDRC.

³⁶ Beatty, B. (2011). The dilemma of scripted instruction: Comparing teacher autonomy, fidelity, and resistance in the Froebelian Kindergarten, Montessori, Direct Instruction, and Success for All. *Teachers' College Record*, 113(3), 395–430.

³⁷ Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, 73(2), 125–230.

³⁸ Enzle, M.E., & Anderson, S.C. (1993). Surveillant intentions and intrinsic motivation. *Journal of Personality and Social Psychology*, 64, 257-266.

³⁹ Ross, J., & Staw, B. M. (1993). Organizational escalation and exit: Lessons from the Shoreham nuclear power plant. *The Academy of Management Journal*, 36(4), 701–732.

⁴⁰ Patil, S., Vieider, F., & Tetlock, P. E. (2012). Process and outcome accountability: *Oxford handbook of public accountability*. New York, NY: Oxford University Press.

⁴¹ Baer, R., Hinkle, S., Smith, K., & Fenton, M. (1980). Reactance as a function of actual versus projected autonomy. *Journal of Personality and Social Psychology*, 35, 416-422.

⁴² Tyack, D., & Hansot, E. (1982). *Managers of virtue: Public school leadership in America, 1820–1980*. New York, NY: Basic Books.

⁴³ Glenn, Jr., C. L. (1988). *The myth of the common school*. Amherst, MA: University of Massachusetts Press.

⁴⁴ Friedman, M. (1955). The role of government in education. In R. A. Solo (Ed.), *Economics and the Public Interest*. Piscataway, NJ: Rutgers University Press.

⁴⁵ Chubb, J., & Moe, T. (1990). *Politics, markets, and America's schools*. Washington, DC: Brookings Institution.

⁴⁶ Coons, J. E. (1992). School choice as simple justice. *First Things*, 22, 193–200.

⁴⁷ Gill, B., Timpane, P. M., Ross, K. E., Brewer, D. J., & Booker, K. (2007). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools*. Santa Monica, CA: The RAND Corporation.

⁴⁸ Abdulkadiroglu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *Quarterly Journal of Economics*, 126(2), 699–748.

⁴⁹ Booker, K., Sass, T. R., Gill, B., & Zimmer, R. (2011). The effects of charter high schools on educational attainment. *Journal of Labor Economics*, 29, 377–415.

⁵⁰ Dobbie, W., & Fryer, R.G. (2015). The medium-term impacts of high-achieving charter schools. *Journal of Political Economy*, 123, 985-1037.

⁵¹ Gleason, P. M., Tuttle, C.C., Gill, B., Nichols-Barrer, I., & Teh, B. Do KIPP schools boost student achievement? *Education Finance and Policy*, 9, 36-58.

⁵² Gill, B., & Nichols-Barrer, I. (2014). Charter schools. In D. Brewer & L. Picus (Eds.), *Encyclopedia of Education Economics and Finance*. Thousand Oaks, CA: Sage Publications, Inc.

⁵³ Center for Research on Education Outcomes (CREDO). (2013). *National charter school study*. Stanford, CA: Stanford University.

⁵⁴ Wolf, P., Kisida, B., Gutmann, B., Puma, M., Eissa, N., & Rizzo, L. (2013). School vouchers and student outcomes: Experimental evidence from Washington, DC. *Journal of Policy Analysis and Management*, 32(2), 246–270.

⁵⁵ Chingos, M. M., & Peterson, P. E. (2015). Experimentally estimated impacts of school vouchers on college enrollment and degree attainment. *Journal of Public Economics*, 122, 1–12.

⁵⁶ Zimmer, R.W., & Toma, E.F. (2000). Peer effects in public and private schools across countries. *Journal of Policy Analysis and Management*, 19(1), 75-92.

-
- ⁵⁷ Bifulco, R., & Ladd, H. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy, 1*(1), 50–90.
- ⁵⁸ Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. *Economics of Education Review, 24*, 133–147.
- ⁵⁹ Zimmer, R., Gill, B., Booker, T. K., Lavertu, S., Sass, T. R., & Witte, J. (2009). *Charter schools in eight states: Effects on achievement, attainment, integration, and competition*. Santa Monica, CA: RAND.
- ⁶⁰ Winters, M. A. (2012). Measuring the competitive effect of charter schools on public school student achievement in an urban environment: Evidence from New York City. *Economics of Education Review, 31*(2), 293–301.
- ⁶¹ Jinnai, Y. (2014). Direct and indirect impact of charter schools' entry on traditional public schools: New evidence from North Carolina. *Economics Letters, 124*, 452–456.
- ⁶² Imberman, S. (2011). The effect of charter schools on achievement and behavior of public-school students. *Journal of Public Economics, 95*, 850–863.
- ⁶³ Chingos, M. M., & Peterson, P. E. (2010). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review, 30*, 449–465.
- ⁶⁴ Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: longitudinal analysis with student fixed effects. *Economics of Education Review, 26*, 673–682.
- ⁶⁵ Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*, 915–945.
- ⁶⁶ Hawley, W., & Valli, L. (1999). The essentials of effective professional development: A new consensus. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151–180). San Francisco, CA: Jossey Bass.
- ⁶⁷ Cowan, J., & Goldhaber, D. (2015). National Board certification and teacher effectiveness: Evidence from Washington. University of Washington Bothell, Center for Education Data and Research.
- ⁶⁸ Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.
- ⁶⁹ Barnwell, P. (2015). Why schools need more “hybrid” teaching roles. *Education Week*, February 18.
- ⁷⁰ Gawande, A. (2011). Personal best. *The New Yorker*, October 3.
- ⁷¹ Newton, L. H., Hodges, L., & Keith, S. (2013). Accountability in the professions: Accountability in journalism. *Journal of Mass Media Ethics, 19*, 166–190.
- ⁷² Porter, A. C., Polikoff, M. S., Goldring, E. B., Murphy, J., Elliott, S. N., & May, H. (2010). Investigating the validity and reliability of the Vanderbilt Assessment of Leadership in Education. *The Elementary School Journal, 111*(2), 282–313.
- ⁷³ Frank, K. (2012). Constitution for effective school governance. Commentary in *Teachers College Record*, February 24, ID number 16715.
- ⁷⁴ Chicago Public Schools. (2014). REACH Students Educator Evaluation Handbook 2014-15.
- ⁷⁵ OECD. (2014). *TALIS 2013 results: An international perspective on teaching and learning*. Paris: OECD Publishing.
- ⁷⁶ Pil, F.K. & Leana, C. (2009). Applying organizational research to school reform: The effects of human and social capital on student performance. *Academy of Management Journal, 52*(6), 1101-1124.
- ⁷⁷ Bryk, A.S. & Schneider, B. (2004). *Trust in schools: a core resource for improvement*. New York: Russell Sage Foundation
- ⁷⁸ Ashton, R.H. (1992). Effects of justification and a mechanical aid on judgment performance. *Organizational Behavior and Human Decision Processes, 52*, 292-306.
- ⁷⁹ Lerner, J.S., Goldberg, J.H., & Tetlock, P.E. (1998). Sober second thought: The effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin, 24*, 563-574.
- ⁸⁰ Bodenhausen, G. V., Kramer, G. P., & Susser, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology, 66*(4), 621–632.
- ⁸¹ Tetlock, P. E. (1985). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly, 48*(3), 227–236.

-
- ⁸² Scholten, L., van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K.W. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology*, 43, 539–552.
- ⁸³ Sherman, G. D., Lee, J. J., Cuddy, A. J. C., Renshon, J., Oveis, C., Gross, J. J., & Lerner, J. S. (2012). Leadership is associated with lower levels of stress. *Proceedings of the National Academy of Sciences*, 109(44), 17903–17907.
- ⁸⁴ Hancock, P.A., & Warm, J.S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, 31, 519-537.
- ⁸⁵ Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.
- ⁸⁶ Tetlock, P. E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. *Advances in Experimental Social Psychology*, 25, 331–376.
- ⁸⁷ Thaler, R.H., & Sunstein, C.B. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- ⁸⁸ Cialdini, R.B. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika*, 72, 263-268.
- ⁸⁹ Hallsworth M., List, J., Metcalfe, R., Vlaev, I. “The Behavioralist as Tax Collector: Increasing Tax Compliance Through Natural Field Experiments,” NBER Working Paper no. 20007, 2014.
- ⁹⁰ Kraft, M., & Rogers, T. (2015). The underutilized potential of teacher-to-parent communication: Evidence from a field experiment. *Economics of Education Review* 47, 49-63.
- ⁹¹ Hallberg, K., & Green, G. (2015, March 11). How can we hire and keep high quality teachers in struggling schools? InformED Blog, American Institutes for Research Education Policy Center.
- ⁹² Furgeson, J., Gill, B., Haimson, J., Killewald, A., McCullough, M., Nichols-Barrer, I., ... Lake, R. (2012). *Charter-school management organizations: Diverse strategies and diverse student impacts*. Cambridge, MA: Mathematica Policy Research.
- ⁹³ Marsh, J. A., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making. *Educational Policy*, 24(6), 872–907.
- ⁹⁴ Blazar, D., & Kraft, M.A. (2015). Exploring mechanisms of effective teacher coaching: A tale of two cohorts from a randomized experiment. *Educational Evaluation and Policy Analysis*, 37(4), 542-566.
- ⁹⁵ Kane, T. J. (2012). Capturing the dimensions of effective teaching. *Education Next*, 12(4).
- ⁹⁶ Chaplin, D., Gill, B., 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, Mid-Atlantic Regional Educational Laboratory.
- ⁹⁷ Walsh, E., & Lipscomb, S. (2013). *Classroom observations from Phase 2 of the Pennsylvania Teacher Evaluation Pilot: Assessing internal consistency, score variation, and relationships with value added*. Cambridge, MA: Mathematica Policy Research.
- ⁹⁸ L’Hommedieu, R., Menges, R. J., & Brinko, K. T. (1990). Methodological explanations for the modest effects of feedback from student ratings. *Journal of Educational Psychology*, 82(2), 232–241.
- ⁹⁹ Taylor, E. S., & Tyler, J. H. (2012). The effect of evaluation on teacher performance. *American Economic Review*, 102(7), 3628–3651.
- ¹⁰⁰ Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2), 267–297.
- ¹⁰¹ Rockoff, J.E., Staiger, D.O., Kane, T.J., & Taylor, E.S. (2012). Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review*, 102(7): 3184-3213.
- ¹⁰² See also Schillemans, T., & Smulders, R. (2015). Learning from accountability? Whether, what, and when. *Public Performance & Management Review*, 39(1): 248-271.
- ¹⁰³ Furgeson, J., McCullough, M., Wolfendale, C., & Gill, B. (2014). *The Equity Project Charter School: Impacts on student achievement*. Cambridge, MA: Mathematica Policy Research.