

## LOOSE COUPLING: THE “CONDITION” AND ITS SOLUTIONS?

WILLIAM A. FIRESTONE

*GRADUATE SCHOOL OF EDUCATION - RUTGERS UNIVERSITY*

**Abstract:** This paper describes conceptualizations of loose coupling developed in the 1970s and uses them to trace the tightening of two coupling mechanisms in the succeeding decades. Advocates for both kinds of tightening saw them as ways to increase student achievement overall and reduce achievement gaps in particular. The first is increase in both research and interventions to tighten cultural couplings in ways that were expected to create more collaborative, learning focused schools for both teachers and students. This approach does seem to contribute to student learning where it happens, but—for all the advertising of programs to create professional learning communities—its application is uneven at best. The second was the increase in (usually testing-based) external accountability, often linked to increased sanctions and rewards. Increasing accountability is more extensive, but its effects appear to be mixed. My own research has contributed to understanding how both these trends have manifested in schools and districts.

Firestone, W. (2015). Loose coupling: The “condition” and its solutions? *Journal of Organizational Theory in Education* 1(1). Retrieved from [www.organizationaltheoryineducation.com](http://www.organizationaltheoryineducation.com).

One concept that had almost a faddish following when I started my career in the '70s was the idea of loosely coupled organizations. The term was popularized by Karl Weick (1976), but came to encompass Meyer and Rowan's (1977) theory of institutionalized organizations, and Lortie's (1969) observations on teacher isolation and autonomy as well. These authors were usually agnostic about whether loose coupling was a problem or a solution. However, as their observations came just when the Coleman Report (Coleman, Campbell, Hobson, McPartland, Mood, Weinfield, & York, 1966) was beginning to publicize the idea that American education was not contributing sufficiently to the well-being of the nation, loose coupling became associated with two problems in American education. First, education did not contribute sufficiently to a productive economy because—as *A Nation at Risk* publicized (National Commission on Excellence in Education, 1983)—American school children were not achieving at the highest levels reached in other countries. Second, the American educational system was not reducing the gross inequities among ethnic groups and between rich and poor.

Some organizational (and policy) research and reform writing in the intervening 40 years has unpacked the loose coupling metaphor. Even more has assumed that loose coupling was a problem to be solved. Two solutions have developed. Organizational researchers, researchers on teaching, consultants and professional educators emphasized tightening cultural couplings to create in-school “professional learning communities” or other arrangements through which local educators could collectively enhance their capacity to teach students. Policy researchers—especially those with an economics bent—and advocates preferred to combine the techniques of academic testing with the authority of the state through the accountability movement. The common thread in this approach has been to measure desired outcomes and practices and provide rewards and punishments contingent on measured results. While both approaches have become very popular, it is doubtful that either have changed teaching very much.

My work has explored both of these forms of

coupling over the last 40 years. In this article, I will briefly review some of the central ideas about loose coupling and then describe what has been learned from research on these very different efforts to tighten couplings in American education. I will then review some work I have done to clarify our understanding in these areas and then quickly summarize what I think has changed and what has not in the educational sector.

## Loose Coupling

The common thread in loose coupling literature was not just that the rational, top-down, authority-driven approaches to understanding school organization that implicitly dominated thinking about school improvement in the '60s and '70s were misguided. The rational explanations made little sense to people who worked in schools because neither the technical core of schooling nor authority with its goals, sanctions, and formal offices and right of command explained much of the action in schools (Weick, 1976). Loose coupling was a sensitizing concept that could be instantiated in several ways, including sparse networks that slowed the spread of information and influence in the system, a situation where different actions always led to similar results, or the intentional delegation of discretion to lower levels of the hierarchy.

Loose coupling was dysfunctional in some ways, but helpful in others. Weick (1976) emphasized the functionality of loose coupling. For instance, loosely coupled organizations can persist in spite of small to modest external disruptions, including most state and federal policy directives. Internally, a breakdown in one part—e.g. one classroom—will not impede the work of the whole school as the malfunction of one part of an assembly line might. Similarly, because loose coupling facilitates local adaptation, it allows individual teachers to better respond to the educational needs of a particular classroom or even specific students than a more tightly coupled entity might. Moreover, loose coupling can be motivational by giving teachers more responsibility and autonomy.

Lortie's (1975) study of school teachers fleshed out what it was like to work in a loosely coupled organization. Teachers' most powerful rewards in the '60s when his data were collected were not the extrinsic status or remuneration from the district but the psychic reward of "knowing that I have 'reached' students and they have learned." Obtaining this reward was prob-

lematic though. Teachers felt alone to work out how to teach their children. They had to develop their own instructional goals which were as much moral and about learning behavior as they were focused on learning the formal curriculum. Moreover, teachers were always uncertain about how their students were doing and how much student progress was attributable to their work.

Meyer and Rowan (1977) looked outside the school system to explain the spread and persistence of loose coupling. They suggested that as society modernized, institutionalized "myths" developed about what a rational organization should be. Organizational structures—i.e., myths—were adopted more for how they looked than for their contribution to productivity. These myths were reinforced by coercive or legal isomorphism where laws required compliance, normative isomorphism where professions imposed order on a field through licensure (often working with the state), and mimetic isomorphism where later adopters copied earlier ones (DiMaggio & Powell, 1983), all to increase the legitimacy of the enterprise in the eyes of a larger public (Meyer & Rowan, 1977). The spread of the superintendent—a role with considerable prestige and authority but limited impact on teaching—as a reflection of the myth of the supremacy of the corporate CEO in early to middle 20th century American society exemplifies mimetic isomorphism (Tyack, 1974). The spread of many specialized positions, from counselors to Title I staff, reflected both normative and coercive processes (Meyer, Scott, & Strang, 1994; Rowan, 1982).

All together, the research of this era portrayed American education as having an organizational structure that reflected external (mostly local, but increasingly state and national) demands but that did not strongly influence teaching and learning. Teachers had the joys and frustrations of coping with the challenges of their work by themselves. To the extent that policy makers and reformers were aware of the loose coupling idea, they saw it as a problem. Prescriptions to make schools more equitable and productive usually entailed tightening cultural or structural couplings.

## Tightening Culture through Professional Community

The idea of increasing coupling through building a professional community or a shared culture has been popular among organizational researchers and researchers studying teaching as well as several types of reformers.

Research on the benefits of tightening cultural coupling has been conducted to advance knowledge about teachers' workplace, school professional climates, change management, professional development, and leadership.

The earliest studies of teachers' workplace drew on Lortie's (1975) analysis of teacher isolation and autonomy, but where Lortie emphasized uniformity, they examined variation among schools. In the '80s researchers found that in schools where teachers shared goals that focused on student learning, had norms and expectations that sustained collegial analysis of teaching, positive developments followed, including more active collaboration among teachers, greater teacher commitment to their work, and more continuous learning about how to teach. These developments could contribute to improved student learning, even after controlling for student characteristics (Little, 1982; Rosenholtz, 1989).

Climate studies have a long history of surveying teachers to find shared perceptions of what a school is like. Recently, researchers have identified climate dimensions that may provide cultural couplings that enhance student learning. One of these is collective efficacy, a belief shared among teachers that they have the capacity to improve the performance of their school (Goddard, Hoy, & Woolfolk Hoy, 2000). Another is trust or the willingness of one party to be vulnerable to another under certain conditions (Tschannen-Moran & Hoy, 2000). When these factors characterize whole schools, they contribute to student learning, again when controlling for student background factors (Bryk & Schneider, 2002; Goddard, Hoy, & Woolfolk Hoy, 2000).

Over time, much more developed conceptions of a collaborative workplace developed. In several studies, Louis and colleagues defined a professional community as a school where teachers shared norms and values about their work, had a collective focus that gave primacy to student learning, had norms that supported collaboration, actually deprivatized practice so teachers saw each other's work in various forms, and engaged in reflective dialogue with in-depth conversation about their teaching (Bryk, Camburn, & Louis, 1999; Louis, Kruse, & Marks, 1996). Some conditions that contributed to such a professional community were sufficient scheduled planning time, a certain level of teacher empowerment, and a climate of trust among educators in the school. Schools with strong professional communities could support more authentic, complex instruction and higher levels of authentic student achievement, again controlling for student background factors (Louis

& Marks, 1998). More recently, studies have combined the teacher workplace and climate approaches and enhanced analysis through the use of tools like network analysis to show the conditions that promote collaboration among teachers and how collaboration can facilitate professional learning and ultimately student learning (Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Daly, Der-Martirosian, Ong-Dean, Park, & Wishard-Guerra, 2011).

The workplace and climate studies identify the tighter couplings in the forms of beliefs and collaborative interactions that can facilitate student learning. However, they provide limited guidance on how to create those conditions. Here the studies of implementation, professional development, and leadership are especially helpful. The most significant, lasting finding of this research is that, at least in the case of grant-based interventions, local context trumps policy. State and federal policies are interpreted locally in very diverse ways (Berman & McLaughlin, 1977).

As implementation began to be conceptualized as at least partly a learning process, collaboration among educators became more central. By the mid-2000s, considerable evidence had been amassed that professional development was enhanced when teachers could work together and use reflective dialogue to turn general principles into concrete plans for one's own classroom (Borko, 2004; Garet, Porter, Desimone, Birman, & Yoon, 2001). Research we conducted at the Consortium for Policy Research in Education (CPRE) showed how networks could be constructed across school boundaries to provide teachers with opportunities for collegial interaction that would promote the same kind of learning that happened in highly collaborative schools (Firestone & Pennell, 1997).

But how would such collaborative, learning-focused cultures be created, especially in contexts that discouraged working together to promote achievement? Beginning with the effective schools studies (Edmonds, 1979), attention has increasingly focused on school and district leadership which is now seen as one of the more important internal factors contributing to student learning (Waters, Marzano, & McNulty, 2003). Recent syntheses suggest that leaders that contribute to student learning set a direction for their schools that prioritizes improved student learning, promote learning to teach better, redesign and manage the organization to facilitate teaching and learning, and create a collaborative, student-learning focused culture (Hallinger, 2005; Leithwood & Sun, 2012; Robinson, Lloyd, & Rowe,

2008). Effective principals appear to share certain leadership responsibilities with their in-school colleagues, create opportunities for collaborative learning, and keep everyone focused on improving achievement (Wahlstrom, Louis, Leithwood, & Anderson, 2010). There have been fewer studies of the microbehaviors of principals to examine how they work with others to accomplish these tasks, but there has even been work in that direction (Firestone & Wilson, 1985; Spillane, Camburn, & Pareja, 2009).

While a great deal of evidence shows that tightening couplings through building shared goals and collegial cultures, the approach also has its critics. Rick Hess (2013) suggests that while this approach is helpful as far as it goes, some schools lack the talent, time, and changed or reinterpreted regulations to significantly improve. Then tightening culture is necessary but not sufficient. Another approach to tightening couplings in the name of reform has actually received much more attention outside the research and professional education communities.

### **Increasing Institutional Pressures and the Rise of Accountability**

Tightening coupling through external accountability has been more obvious than culture-building and collaboration if less well received by educational professionals. As if responding to the popular critique of loose coupling as well as elite demands for greater productivity, the accountability movement centralized authority over education, not at the school or district level but in state and federal government. This centralization was facilitated by developments in the technology of testing that made large-scale measurement of student learning more feasible and accurate (Herman & Baker, 2009). The spread of policies that first held schools and then teachers accountable with increasingly strong sanctions for poor performance reflects a mix of both mimetic and coercive isomorphism (DiMaggio & Powell, 1983).

The accountability movement began with states' adoption of minimum competency testing in the 1970s. The rapid spread of this policy was an explicit response to public concern that schools were insufficiently productive (Resnick, 1980). Over the next 20 years, the accountability movement became increasingly complex. States adopted progressively more rigorous tests. An important related development was the push for state and national learning standards. During the 1990s—led by the National Council of Teachers of Mathematics

(National Council of Teachers of Mathematics, (NCTM), 1989) and others—states began adopting curriculum or outcome standards and aligning their assessments, curricula and other policies with those standards. Sanctions also became more extensive for educators. Originally, sanctions were applied largely to students who had to pass these tests to be promoted to a higher grade or graduate from high school, actions that only affected educators indirectly. In time, a few states began to put low performing schools on probation and to link accreditation and funding to test scores (Goertz, Duffy, & Consortium for Policy Research in Education, 2001).

Through the '80s, increasing testing requirements was largely a mimetic policy with organizations like the Council of Chief State School Officers and the Education Commission of the States sharing information about what different states were doing among their members. In the late 1980s, through an educational summit and the National Educational Goals Panel, the federal government became involved. Initially, its role was facilitative. Although efforts to develop national standards and voluntary national tests initially failed, this work promoted substantial development and sharing of ideas.

In the '90s, the federal policy became more coercive. First in 1994, and then much more intensively in 2001 with the passage of No Child Left Behind (NCLB), Congress began using the federal Elementary and Secondary Education Act (ESEA) to enhance accountability by specifying criteria for which educators would be held accountable and creating sanctions for non-compliance. NCLB required that to receive ESEA funding, states had to test students in math and language arts in every grade from 3 to 8 and once in high school. Test scores had to be disaggregated to track achievement among groups with a history of poorer school performance. States had to take remedial action for schools that did not make “adequate yearly progress.” These actions could eventually include reconstituting a school, changing its staff, or closing it.

The next step would be to narrow the focus of accountability from the whole school to the individual teacher. Some analysts argued that this narrowing of focus would be facilitated by the development of value-added assessment which, it was argued, allowed one not only to determine the general contribution that teaching made to student learning but to attribute the growth of specific classes of students to their particular teachers

(Wright, Horn, & Sanders, 1997). While this assertion has been challenged (Baker, Barton, Darling-Hammond, Haertel, Ladd, Linn, & Shepard, 2010), it fostered substantial development in the use of student testing, direct observation, and other measures alone and in combination to assess teacher quality (Bill and Melinda Gates Foundation, 2013).

Some experimenting with linking teacher evaluation to various high stakes—including both financial incentives and removing incompetent teachers—has been mimetic. Districts like Washington, DC, Chicago, Charlotte-Mecklenberg, and Hillsborough County, FL have experimented with new forms of teacher evaluation. Some—including Denver, Pittsburgh, and Newark—have linked incentives to value-added test scores or other evaluation results. The spread of teacher evaluation was facilitated by two federal policies that provided incentives for adopting teacher evaluation practices: the Teacher Incentive Fund and the even higher impact Race to the Top (RTTT) competition that gave successful states development funds if they promised to adopt such practices as teacher evaluation based partly on student growth as measured by test scores (Heyburn, Lewis, & Ritter, 2010; US Department of Education, 2009). Although only 18 states received RTTT funds, 35 states now require that student data be used to evaluate teachers (Sawchuck, 2013).

Whether tightening coupling through accountability policies would improve productivity, as its advocates hoped, would depend in part on what kind of loose coupling exists in schools. If it is insufficiently centralized authority, poor measurement systems and inadequate sanctions as critics argue, then tightening coupling by increasing authority might increase student learning and increase equity of outcomes. If it is a functional response to the diverse and unpredictable problems of teaching students as technological determinists suppose or the need to add structures to demonstrate compliance with external demands while protecting teaching from contradictory, unrealistic expectations as institutional theorists propose, accountability policies could be dysfunctional. In between, dysfunctions could also occur if appropriately tightened couplings could facilitate outcomes but poorly designed ones made things worse.

Over the last 25 years the effects of state testing policies have become reasonably clear in summaries by assessment researchers (Hamilton, Stecher, & Yuan, 2008). In some regards, high stakes testing has gotten

educators' attention. They give more evidence to tested subjects than untested ones, focus on students who risk not passing tests—rather than those who clearly will or will not pass those tests (Booher-Jennings, 2005)—revise teaching materials and activities to look more like test items, and in extreme cases cheat to make sure that students and schools pass the test (Levitt & Dubner, 2005). However, there are limits to what testing and sanctions can accomplish. Even in the face of increased accountability, teachers cannot change their teaching in ways they do not understand (Firestone, Mayrowetz, & Fairman, 1998).

Increased accountability has increased student achievement on some measures, but the meaning of that increase is not clear. Scores on the tests used for accountability purposes often go up and do so more when the stakes are higher. Yet, those gains do not always transfer to other tests, most notably NAEP, suggesting that the learning that is measured has quite a bit to do with test-specific knowledge rather than increases in students broader understanding of the subjects taught (Hamilton, Stecher, & Yuan, 2008).

The effects of more focused teacher accountability policies aren't so well known yet. The pros and cons of different ways of measuring teacher effectiveness—especially value-added assessment of individual teachers (Baker, Barton, Darling-Hammond, Haertel, Ladd, Linn, & Shepard, 2010; Sanders & Horn, 1994) are still being debated as are various ways to combine different measurement strategies (Bill and Melinda Gates Foundation, 2013). Financial incentives linked to measured student outcomes may increase teacher motivation when goals are clear, incentives are big enough to matter, and the program gives the teachers clear feedback (Kelley, Odden, Milanowski, Heneman, & Herbert, 2000). However, some of the best designed studies of teacher incentive programs indicate that they have little if any effect on teachers or students (Yuan, Vi-Nhuan, McCaffrey, Marsh, Hamilton, Stecher, & Springer, 2012). However, for all the well-designed experimental and quantitative studies of these programs, the careful qualitative analysis that provided so much insight into the processes surrounding test-based accountability have so far been less in evidence in research on teacher evaluation.

## **The Direction of My Research**

Although my research has taken advantage of many opportunities to address the policy issues of the time, a

consistent theme has been that couplings that promote professional community, intrinsic incentives, and teacher learning are more constructive than those that rely on authority and extrinsic incentives to promote compliance. In 1973 my first job, with a private research company put me in the middle of the newly developing study of educational change. I spent three years working for a major evaluation company and living in a small rural school district that had received a very large federal grant. The result of that work was a detailed case study of how a superintendent of schools reached out to a major federal agency to obtain funding for his innovation of choice and how he was unable to implement that innovation partly because of limits to his authority and internal disagreements. The problem was not so much resistance to change--a popular idea at the time--as limited cultural and authority-based coupling combined with ineffective planning (Firestone, 1980).

On one of my occasional trips to headquarters, I stumbled on Karl Weick's article on loose coupling and was fascinated. A few years later, with my colleagues Bob Herriott and Bruce Wilson, I began a study of loose coupling. While other researchers emphasized the overall loose coupling of schools when compared to other kinds of organization, we focused on variation among schools. Our work showed that on an array of bureaucratic coupling mechanisms, elementary schools were more tightly coupled than high schools. Elementary teachers agreed more on what the school's goals should be than did high school teachers (Herriott & Firestone, 1984). Elementary teachers also agreed more with their principal on what the school's goals than did high school teachers (Wilson, Herriott, & Firestone, 1991). Finally, while teachers always reported having more influence over instructional decisions than their principals, the difference favored teachers substantially more in high schools. In fact, schools consistently grouped into a relatively tightly coupled high goal-consensus, more centralized authority cluster and a loosely coupled one where goal consensus and centralization were both lower; elementary schools were almost uniformly in the first cluster and high schools were in the second. Middle schools tended to be in the loosely coupled cluster but not as consistently as high schools (Herriott & Firestone, 1984). When we looked for organizational or staffing characteristics that might explain this variation in coupling patterns, we found that high schools were larger than elementary schools, more departmentalized, and had more consistently male staff. However, while these differences were correlated with differences in coupling, they did not explain those

differences (Wilson, Herriott, & Firestone, 1991) which added evidence to the arguments that either institutional factors or the technology of teaching determined the patterns of organizational coupling in schools.

Still later, I moved to Rutgers University and began a fruitful eight years of participation in CPRE. This provided a useful opportunity to explore different mechanisms that were being recommended to policy makers to tighten coupling in order to reform education. I had a birds-eye view of the spread of accountability policy in the late '80s and early '90s (Firestone, Fuhrman, & Kirst, 1989; Firestone, 1990), but I also explored the loosely coupled nature of policy making by applying the metaphor of an ecology of separate games to the interaction of various actors (Firestone, 1989). At CPRE, I also explored an earlier interest in how different coupling mechanisms resulting from organizational structures and state policies affected teacher commitment (Firestone & Rosenblum, 1988). This work suggested that formal arrangements like merit pay that stressed extrinsic sanctions did less to build teacher commitment than changes like career ladders that promoted collegiality and professional dialogue (Firestone, 1991).

Somewhat later, I examined how state-testing policies affected teaching in the pre-NCLB era. The themes of this work have been that tightening coupling by changing assessments can modify the surface features of teaching but that deeper changes require transforming teachers' understanding of the content and processes of instruction. Moreover, as suggested by the loose coupling and implementation literature, the same policy can be defined quite differently in different locations.

Maine and especially Maryland in the 1990s, had adopted performance-based assessments in mathematics that should have encouraged teaching for deeper understanding. We learned from our intensive interviews and classroom observations that teachers were influenced by the format of the state tests to teach mathematics using long projects with more complex than usual representations to make connections to the real world. Nevertheless, conventional approaches to instruction where teachers would tell students what procedures to use and how to carry them out and then have students practice those procedures continued to dominate teaching because that was how teachers knew to teach (Firestone, Mayrowetz, & Fairman, 1998). This pattern of affecting surface patterns of teaching

more than deep structure even applied internationally, as I found when I had a chance to conduct similar field work in Wales where the national government required the use of assessments like American performance assessments (Firestone, Broadfoot, & Fitz, 1999).

A larger project in New Jersey used statewide teacher surveys to complement classroom observation. We also collected more data on how district administrators interpreted this state testing program that combined short open-ended and more conventional multiple-choice assessments and explored the local contextual factors that affected how educators interpreted state policy. In this state, the closest thing to sanctions was pressure that came largely from school and district administrators who wanted publicized test scores to make the district look good. Generally, we found that administrative pressure encouraged didactic, drill-oriented math and science instruction and focused test preparation. More inquiry-oriented instruction that encouraged deeper understanding of those subjects was supported by access to learning opportunities, administrative support, and the manipulatives and calculators to facilitate deeper inquiry. This study highlighted the role of district administrators in particular as interpreters of state policy. Where they believed the new testing program and related reforms provided a chance to promote deeper learning, resources—including learning opportunities and appropriate materials—were likely to be made available and more inquiry-oriented instruction followed. Where they saw these policies as a compliance problem, pressure and test-prep oriented professional development dominated (Firestone, Schorr, & Monfils, 2004).

Two later projects examined how districts orchestrated a variety of coupling mechanisms to create learning opportunities for teachers (Firestone, Mangin, Martinez, & Polovsky, 2005; Firestone & Martinez, 2007). Some of these were cultural, like having a shared vision of what constituted good teaching. However, even cultural couplings might be enacted through formal authority. For instance, a district might demonstrate its vision of good teaching by demanding that all schools adopt one particular whole school reform program rather than giving schools a choice as the state, which required schools to choose such programs, had expected. The district would further demonstrate that vision by providing time and curriculum materials to support that vision, and by monitoring teachers' instruction to ensure compliance with that vision, and by supporting that vision through the full array of profes-

sional development provided. The distinction between cultural coupling and bureaucratic ones was blurred as leaders used the latter to reinforce the former, a pattern we had anticipated years ago (Firestone & Wilson, 1985).

Most recently, I have followed trends in state policy by exploring the local effects of using teacher evaluation to tighten couplings. We just completed an assessment of the implementation of New Jersey's pilot teacher evaluation program which suggested some challenges of implementing this program—e.g., finding time to conduct all observations required by state teacher evaluation requirements—and some strengths and weaknesses of this program as a motivator and learning opportunity for teachers. For instance, the classroom observation and student growth data provided by the program helps teachers to reflect on their teaching. However, even though most teachers do not see the program as a threat to their job stability, the negative incentives involved—along with the timing of delivery of the data and questions about its accuracy—threaten its value as a learning opportunity (Firestone, Blitz, Gitomer, Gradinarova, Kirova, Shcherbakov, & Nordin, 2013). We are currently designing further research to explore how both teachers and administrators understand and use teacher evaluation data.

### **Reflection on 40 Years of Research on Coupling in Education**

In the last 40 years, real changes have taken place in how education is organized. Clearly the accountability movement has had a major impact. The language of standards is much more extensive than it used to be although its most obvious and influential manifestation is in testing. Student testing by central authorities has expanded dramatically and drives decisions about the fates of schools, educators, and students. Teacher evaluation is becoming much more formalized in adopting the trappings of educational measurement.

The language of professional community is also much more widespread. One need only examine the advertisements for training on “professional learning community” in any issue of *Education Week* to see the extensive infrastructure of consultants and trainers that has developed. The idea of professional community has even been mandated in at least one state through New Jersey's professional standards for teachers (New Jersey Department of Education, 2004).

Yet, in keeping with institutional theory, structural responses seem more widespread than changes in practice. We do have more testing, labeling (schools “in need of improvement”), and accountability (takeover of failing schools). We do have more trainers ready to help educators develop a professional learning community. Yet, responses to accountability policy are limited to what can be most easily regulated—e.g., content taught but not instructional methods used. The limits to regulation seem to come through the extent of teachers’ knowledge and capacity to respond to the learning needs of particular children. Professional community can provide an opportunity to develop that knowledge, but only where leaders fully understand and embrace the idea and adjust the formal structures to make it possible. Thus, although education is more tightly coupled than it was 40 years ago, it doesn’t seem to be more effective, except in pockets where teachers and their leaders know how to integrate different coupling mechanisms to create a more productive learning environment. What to do about that is a puzzle that will keep organizational theorists and other educational researchers busy for a long time.

## References

- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute.
- Berman, P., & McLaughlin, M. W. (1977). *Federal programs supporting educational change, v. VII, factors affecting implementation and continuation*. Santa Monica: RAND.
- Bill and Melinda Gates Foundation. (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project’s three-year study*. Seattle, WA: Author.
- Booher-Jennings, J. (2005). Below the bubble: “educational triage” and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(3), 3-15.
- Bryk, A. S., Camburn, E., & Louis, K. S. (1999). Professional community in Chicago elementary schools: Facilitating factors and organizational consequences. *Educational Administration Quarterly*, 35(Supplement), 751-781.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation.
- Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. Q. (2010). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.
- Coleman, J. S., Campbell, E. C., Hobson, C., McPartland, J. M., Mood, J., Weinfield, F., & York, (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Daly, A. J., Der-Martirosian, C., Ong-Dean, C., Park, V., & Wishard-Guerra, A. (2011). Leading under sanction: Principals’ perceptions of threat rigidity, efficacy, and leadership in underperforming schools. *Leadership and Policy in Schools*, 10, 171-206.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Isomorphism, and collective rationality in organizational fields. *American Journal of Sociology*, 48, 147-160.
- Edmonds, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Firestone, W. A. (1980). *Great expectations for small schools: The limitations of federal projects*. New York, NY: Praeger.
- Firestone, W. A., Blitz, C. L., Gitomer, D. H., Gradinarova-Kirova, D., Shcherbakov, A., & Nordin, T. L. (2013). *Year 1 report: New Jersey teacher evaluation pilot program*. New Brunswick, NJ: Rutgers Graduate School of Education.
- Firestone, W. A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis*, 11(2), 151-165.
- Firestone, W. A., Mangin, M. M., Martinez, M. C., & Polovsky, T. (2005). Leading coherent professional development: A comparison of three districts. *Educational Administration Quarterly*, 41(3), 413-448.
- Firestone, W. A., Schorr, R. Y., & Monfils, L. (2004). *The ambiguity of teaching to the test*. Mahwah, NJ: Lawrence Erlbaum and Associates.
- Firestone, W. A. (1989). Educational policy as an ecology of games. *Educational Researcher*, 18(7), 18-24.
- Firestone, W. A. (1990). Continuity and incrementalism after all. In J. Murphy (Ed.), *The educational reform movement of the 1980s: Perspectives and cases*. (pp. 143-266). Berkeley, CA: McCutchan.

- Firestone, W. A. (1991). Merit pay and job enlargement as reforms: Incentives, implementation, and teacher response. *Educational Evaluation and Policy Analysis*, 13(3), 269-288.
- Firestone, W. A., Broadfoot, P., & Fitz, J. (1999). Power, learning, and legitimation: Assessment implementation across levels in the US and the UK. *American Educational Research Journal*, 36(4), 759-796.
- Firestone, W. A., Fuhrman, S. H., & Kirst, M. (1989). *The progress of reform: An appraisal of state education initiatives*. CPRE research report series RR-014. New Brunswick, NJ: Center for Policy Research in Education.
- Firestone, W. A., & Martinez, M. C. (2007). Districts, teacher leaders, and distributed leadership: Changing instructional practice. *Leadership and Policy in Schools*, 6(1), 3-35.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998a). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Firestone, W. A., & Pennell, J. R. (1997). Designing state-sponsored teacher networks: A comparison of two cases. *American Educational Research Journal*, 34(2), 237-266.
- Firestone, W. A., & Rosenblum, S. (1988). Building commitment in urban high schools. *Educational Evaluation and Policy Analysis*, 23(4), 285-300.
- Firestone, W. A., & Wilson, B. (1985). Using bureaucratic and cultural linkages to improve instruction: The principal's role. *Educational Administration Quarterly*, 21(2), 7-30.
- 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(4), 915-946.
- Goddard, R. D., Hoy, W. K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measurement, and impact on student achievement. *American Educational Research Journal*, 37(2), 479-507.
- Goertz, M. E., Duffy, M. C., & Consortium for Policy Research in Education. (2001). *Assessment and accountability systems in the 50 states, 1999-2000*. CPRE research report series.
- Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools*, 4(3), 221-239.
- Hamilton, L. S., Stecher, B. M., & Yuan, K. (2008). *Standards-based reform in the United States: History, re-research, and future directions*. Santa Monica, CA: Center on Education Policy. RAND Corporation.
- Herman, J. L., & Baker, E. L. (2009). Assessment policy: Making sense of the babel. *Handbook of education policy research* (pp. 176-190). New York, New York: Routledge.
- Herriott, R. E., & Firestone, W. A. (1984). Two images of schools as organizations: A refinement and extension. *Educational Administration Quarterly*, 20(4), 41-57.
- Hess, F. W. (2013). *Cage-busting leadership*. Cambridge, MA: Harvard Education Press.
- Heyburn, S., Lewis, J., & Ritter, G. (2010). *Teacher incentive fund grantees*. Nashville, TN: National Center on Performance Incentives, Vanderbilt University.
- Kelley, C., Odden, A., Milanowski, A., Heneman, H. I., & Consortium for Policy Research in Education. (2000). *The motivational effects of school-based performance awards*. CPRE policy briefs.
- Leithwood, K., & Sun, J. (2012). The nature and effects of transformational school leadership: A meta-analytic review of unpublished research. *Educational Administration Quarterly*, 48(3), 387-423.
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. New York, NY: Harper-Collins.
- Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, 19(3), 325-340.
- Lortie, D. C. (1969). The balance of control and autonomy in elementary school teaching. In A. Etzioni (Ed.), *The semi-professions and their organization*. New York, NY: Free Press.
- Lortie, D. C. (1975). *Schoolteacher: A sociological analysis*. Chicago, IL: University of Chicago Press.
- Louis, K. S., Kruse, S. D., & Marks, H. M. (1996). Schoolwide professional community. In F. M. Newmann (Ed.), *Authentic achievement: Restructuring schools for intellectual quality* (pp. 179-203). San Francisco, CA: Jossey-Bass.
- Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 106, 534-575.
- Meyer, J. H., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83, 340-363.

- Meyer, J. H., Scott, W. R., & Strang, D. (1994). Centralization, fragmentation and school district complexity. In W. R. Scott, & J. H. Meyer (Eds.), *Institutional environments and organizations: Structural complexity and individualism* (pp. 160-178). Thousand Oaks, CA: Sage.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: US Government Printing Office.
- National Council of Teachers of Mathematics, (NCTM) (1989). *Curriculum and evaluation standards for school mathematics*. Reston: VA: NCTM.
- New Jersey Department of Education. (2004). *New Jersey professional standards for teachers and school leaders*. Trenton, NJ: New Jersey Department of Education.
- Resnick, D. P. (1980). Minimum competency testing historically considered. In D. C. Berliner (Ed.), *Review of research in education* (Vol 8 ed., pp. 3-29). Washington, DC: American Educational Research Association.
- Robinson, V. M. J., Lloyd, C., & Rowe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44(5), 635-674.
- Rosenholtz, S. J. (1989). *Teachers' workplace: The social organization of schooling*. New York, NY: Longman Inc.
- Rowan, B. (1982). Organizational structure and the institutional environment: The case of public schools. *Administrative Science Quarterly*, 27(2), 259-279.
- Sanders, W. L., & Horn, S. (1994). The Tennessee value-added assessment system (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8, 299-311.
- Sawchuck, S. (2013, February 6). Teachers' ratings still high despite new measures: Changes to evaluation systems yield only subtle differences. *Education Week*, pp. 1-19.
- Spillane, J. P., Camburn, E., & Pareja, A., S. (2009). School principals at work: A distributed perspective. In K. Leithwood, B. Mascall & T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 87-110). New York, NY: Routledge.
- Tschannen-Moran, M., & Hoy, W. K. (2000). A multi-disciplinary analysis of the nature, meaning, and measurement of trust. *Review of Educational Research*, 70(4), 547-593.
- Tyack, D. B. (1974). *The one best system*. Cambridge, MA: Harvard University Press.
- US Department of Education. (2009). *Race to the top program: Executive summary*. Washington, DC: Author.
- Wahlstrom, K. L., Louis, K. S., Leithwood, K., & Anderson, S. E. (2010). *Learning from leadership project: Investigating the links to improved student learning*. Minneapolis, MN: University of Minnesota, Center for Applied Research and Educational Improvement.
- Waters, J. T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Denver, CO: McREL.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21(1), 1-19.
- Wilson, B. L., Herriott, R. E., & Firestone, W. A. (1991). Explaining differences between elementary and secondary schools: Individual, organizational, and institutional perspectives. In P. W. Thurston, & P. P. Zodhiates (Eds.), *Advances in educational administration* V. 2. (pp. 131-157). Greenwich, CT: JAI Press.
- Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.
- Yuan, K., Vi-Nhuan, L., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2012). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation & Policy Analysis, online version*  
doi:10.3102/016237371246262