

DETAILS, DETAILS, DETAILS

RICHARD F. ELMORE*

Professors Liebman and Sabel have written a landmark piece of legal, political, and institutional analysis about the landscape of school reform over the past thirty or so years.¹ It is in the best tradition of scholarly legal analysis, focusing on the implications of judicial and legislative decisions in the context of complex social and institutional forces. I find few points of disagreement on the specifics of the argument Liebman and Sabel present, but I am skeptical about the overall tone of the analysis. If we accept their view of school reform, writ large, it is a highly teleological process—moving toward a higher, better end—or, at the very least, a Hegelian process—moving through alternate stages of opposition and synthesis toward a higher, better end. As a student of the implementation of public policy over the last thirty years, I subscribe to the “anything that can go wrong, will” school of thought, and, on my better days, to the “devil is in the details” school. So while generally affirming Liebman and Sabel’s conclusions, I take friendly exception to their optimistic tone.

For the past several years, my colleagues and I have been doing research on schools’ responses to accountability systems and on the processes of improvement that schools and school systems go through when they increase student performance. Some of this work is summarized—accurately, I might add—in the Liebman and Sabel article. Some is not. In general, my conclusions about the “new accountability” in public education are parallel with those of Liebman and Sabel. Performance-based accountability systems seem to be associated with aggregate increases in student performance. The stronger and tighter the accountability requirements, the greater the improvements, and the more those improvements seem to be associated with traditionally low-performing student populations. Secondary schools have huge systemic problems of student retention and dropping out, but our work suggests that the introduction of so-called “high stakes accountability” has not aggravated these problems, and, in fact, may have raised their visibility. In addition, we find, as Liebman and Sabel suggest, that internal accountability is an important predictor and determinant of how effectively schools respond to the demands of external accountability systems. That is, schools that manifest clarity of mission and purpose, agreement on high

* Gregory Anrig Professor of Educational Leadership at the Graduate School of Education, Harvard University, and Senior Research Fellow at the Consortium for Policy Research in Education.

1. James S. Liebman & Charles F. Sabel, *A Public Laboratory Dewey Barely Imagined: The Emerging Model of School Governance and Legal Reform*, 28 N.Y.U. REV. L. & SOC. CHANGE 183 (2003).

quality instructional practice, and acceptance of responsibility for student performance, individually and collectively, seem to be able to make decisions consistent with success under new accountability systems. Likewise, schools with low internal accountability seem not to be able to work effectively in accountability systems.²

With these findings as background, let me suggest some reasons why I do not fully share the teleological optimism of the Liebman and Sabel piece.

TESTING AND CAPACITY-BUILDING

Policies are bundles of instruments, each with its own political logic, its own technical characteristics, and its own theory of action. In accountability policies there are three primary instruments: standards, testing and investment in human knowledge and skill, or capacity-building. An important feature of the politics of accountability that Liebman and Sabel do not stress is the tendency of accountability policies, when left to their own devices, to drift toward emphasis on testing as the primary instrument, and to de-emphasize standards and capacity-building. The reasons for this are clear: testing is relatively cheap; the less sophisticated the test, the cheaper it is. Standards open up political battles and curriculum wars that most public officials find tedious and risky. Investments in educators' knowledge and skill seem to most public officials to be special-interest policies favoring organized professional interests that do not carry broad constituency appeal.

As I have suggested in another article, cited by Liebman and Sabel,³ I think the No Child Left Behind Act of 2001⁴ (NCLB) embodies a very dangerous drift toward testing as the primary instrument of accountability policy. The law imposes a single regime of annual yearly progress on all schools and school systems, it requires testing in all grades between grades three and eight, and it imposes annual performance increments that carry strong incentives either for states with high-level assessment systems to dumb down their tests or states with already low-level tests to make them lower. In other words, the law rewards schools and school systems essentially for gaming the test, rather than for setting high and challenging standards and using testing *and* human investment together as strategies for improving quality and performance.

I am lately spending a lot of time in schools that have been, or soon will become, classified as failing schools under NCLB. Somewhat surprisingly to me, what I am finding is that many of these schools are fully aware of their performance, they are engaging in focused work on curriculum and instructional

2. See THE NEW ACCOUNTABILITY: HIGH SCHOOLS AND HIGH-STAKES TESTING (Martin Carnoy, Richard Elmore & Leslie Siskin eds., 2003).

3. See Liebman & Sabel, *supra* note 1, at n.459 and accompanying text (citing Richard F. Elmore, *Unwarranted Intrusion*, EDUC. NEXT (Spring 2002), at <http://www.educationnext.org/20021/30.html>).

4. Pub. L. No. 107-110, 115 Stat. 1425 (2001).

practice, and, for the most part, they have internalized the main message of performance-based accountability systems that they are individually and collectively responsible for student learning. That's the good news.

The bad news is that these schools will not meet their performance targets under NCLB, not because they are not improving, but because the law is based on a faulty theory that improvement occurs in equal annual increments, while actual improvement moves in a much different pattern. The bad news also is that states and localities are, for the most part, seriously under-investing in the capacity of teachers and administrators to monitor and improve instructional practice.⁵

My point here is that the teleological view that we are evolving toward a grand synthesis of competing views of school reform understates the possibility that we can make, and probably are making, huge strategic errors that undermine the essential purposes of standards-based school reform. Policymakers have shown a willingness to ignore expert advice on these matters. All of the problems that are present in NCLB were accurately predicted and fully defined by a series of studies commissioned by the National Research Council specifically to inform the reauthorization of Title 1.⁶ Sometimes the political logic of reforms undermines their essential purposes.

WHAT HAPPENED TO THE "LABORATORY OF FEDERALISM?"

The argument that Liebman and Sabel make about the convergence of various reform strategies at the state, local, and school levels can be somewhat deceiving. In fact, as many critics of standards-based reform have often argued, it is a potentially risky and untested idea in many of its current forms. It is also true that states and localities were trying various forms of performance-based accountability, based on distinctively different design principles, and these design variations were the source of much learning about the successes and pitfalls of accountability policy. Some states, Massachusetts, Maryland, New York, and Vermont, for example, deliberately chose relatively high-level tests and set relatively long-run targets for schools to meet performance expectations. Other states, Texas, for example, chose relatively low-level tests and standards on the assumption that, over time, they would raise them as schools began to understand the new incentive structures. No state has satisfactorily addressed the

5. Preliminary results of this research are reported in RICHARD F. ELMORE, NAT'L GOVERNORS ASS'N CTR. FOR BEST PRACTICES, *KNOWING THE RIGHT THING TO DO: SCHOOL IMPROVEMENT AND PERFORMANCE-BASED ACCOUNTABILITY* (2003), available at http://www.nga.org/cda/files/0803_KNOWING.PDF.

6. See especially the work of the National Research Council's Board on Testing and Assessment, e.g.: *TESTING, TEACHING, AND LEARNING: A GUIDE FOR STATES AND LOCALITIES* (Richard F. Elmore & Robert Rothman eds., 1999); *HIGH STAKES: TESTING FOR TRACKING, PROMOTION, AND GRADUATION* (Jay P. Heubert & Robert M. Hauser eds., 1998); *EDUCATING ONE AND ALL: STUDENTS WITH DISABILITIES AND STANDARDS-BASED REFORM* (Lorraine M. McDonnell, Margaret J. McLaughlin & Patricia Morison eds., 1997).

capacity issues noted above, but as Liebman and Sabel suggest, there were some promising developments around capacity-building in Texas and Kentucky.

Now, under the requirements of NCLB, the incentives are essentially (a) to lower standards, and (b) to use tests that are easy and inexpensive to administer. NCLB will have a homogenizing effect on school reform in the U.S. because states and localities will be under tremendous pressure to reduce the number of schools potentially classified as failing under the law and to move toward standards and testing regimes that permit measurement in simple annual increments.

There is nothing inherent in the idea of performance-based accountability that requires annual testing. In fact, one can make a defensible argument, using Liebman and Sabel's language of "internal, strategic, or diagnostic" standards and assessments, that high stakes tests should not be given annually, and that schools and systems should be urged to use their own more curriculum-based measures to monitor improvement. High stakes tests could be used at longer intervals. One could also make the argument that states that had high standards, high-level tests, and relatively low stakes for schools and individual students—Vermont, for example—should be allowed to experiment with and develop those systems as alternatives to the dominant design of annual, high-stakes testing.

My point here is that prior to NCLB, we had a policy regime of experimentation in the laboratory of federalism that was consistent with the relatively weak level of knowledge and high level of uncertainty inherent in performance-based accountability. This state of affairs has been a huge strategic asset in the development of the logic of the policy. NCLB narrows the domain of experimentation drastically and hence limits the amount we can know. It may also be the case that, contrary to the logic of Liebman and Sabel's argument, there is no unified theory of school reform toward which we are evolving, but rather a collection of theories, with various instantiations, organized around a common set of themes, and learning—at the individual and collective level—is what the enterprise is about.

The devil is, in fact, in the details. It is possible to make huge and costly errors following perfectly understandable strands of political, legal, and institutional logic. It is possible that, having made those mistakes, the legitimacy and authority of a whole accountability system can be called into question, and what seems a plausible policy idea can be made into a silly, discredited caricature of itself.

=