

# RESEARCH IN SOCIOLOGY OF EDUCATION

Series Editors: Bruce Fuller and Emily Hannum

Previous Volumes:

- Volume 10: Research in Sociology of Education and Socialization – Edited by Aaron M. Pallas
- Volume 11: Research in Sociology of Education and Socialization – Edited by Aaron M. Pallas
- Volume 12: Research in Sociology of Education and Socialization – Edited by Aaron M. Pallas
- Volume 13: Schooling and School Capital in Diverse Communities – Edited by Bruce Fuller and Emily Hannum
- Volume 14: Inequality Across Societies: Families, Schools and Persisting Stratification – Edited by David Baker, Bruce Fuller, Emily Hannum and Regina Werum
- Volume 15: Children's Lives and Schooling Across Societies – Edited by Bruce Fuller and Emily Hannum

RESEARCH IN SOCIOLOGY OF EDUCATION VOLUME 16

**STRONG STATES, WEAK  
SCHOOLS: THE BENEFITS  
AND DILEMMAS  
OF CENTRALIZED  
ACCOUNTABILITY**

EDITED BY

**BRUCE FULLER**

*University of California, Berkeley, USA*

**MELISSA K. HENNE**

*University of California, Berkeley, USA*

**EMILY HANNUM**

*Oxford University, UK*



JAI

United Kingdom – North America – Japan  
India – Malaysia – China

JAI Press is an imprint of Emerald Group Publishing Limited  
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2008

Copyright © 2008 Emerald Group Publishing Limited

**Reprints and permission service**

Contact: [booksandseries@emeraldinsight.com](mailto:booksandseries@emeraldinsight.com)

No part of this book may be reproduced, stored in a retrieval system, transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without either the prior written permission of the publisher or a licence permitting restricted copying issued in the UK by The Copyright Licensing Agency and in the USA by The Copyright Clearance Center. No responsibility is accepted for the accuracy of information contained in the text, illustrations or advertisements. The opinions expressed in these chapters are not necessarily those of the Editor or the publisher.

**British Library Cataloguing in Publication Data**

A catalogue record for this book is available from the British Library

ISBN: 978-1-84663-910-4

ISSN: 1479-3539 (Series)



Certificate number .....1985.....

Awarded in recognition of  
Emerald's production  
department's adherence to  
quality systems and processes  
when preparing scholarly  
journals for print



INVESTOR IN PEOPLE

## ACKNOWLEDGMENTS

We benefited greatly over the past year from a third co-editor, Melissa Henne, a wonderfully talented young scholar. Her keen knowledge of social theory and the state's contemporary effort to exact local accountability served to advance the sophistication of the chapters that appear in this volume.

Tough-headed reviewers also help to ensure quality results. Again this year we counted on the editorial board, including Fran Vavrus, Annette Lareau, Sara Lawrence-Lightfoot, Aaron Pallas, Chiqui Ramirez, Steve Raudenbush, and Yoshi Shavit. In addition, Cynthia Coburn, Michal Kurlaender, Judith Warren Little, and Jennifer Russell provided thoughtful reviews. Sitome Mebrahtu aided the editing process and injected a strong dose of her collective spirit.

*Research in Sociology of Education* remains committed to publishing the best work that advances our theoretical accounts of important educational problems and pushes forward on innovative empirical methods. You will notice that chapters appearing in the series are crafted by junior and senior scholars in sociology, economics, political science, and anthropology. Thank you all for contributing to this sixteenth edition.

Bruce Fuller  
Emily Hannum  
*Series Editors*

# CHAPTER 1

## OVERVIEW: LIBERAL LEARNING IN CENTRALIZING STATES

Bruce Fuller

What I needed from the classroom was a public life ... not individuality but something closer to the reverse. Teach me about mad British kings so I will understand the American penchant for iconoclasm. The classroom will touch a history that implicates us with others ... that we belong to a culture.

– Richard Rodriguez, *Days of Obligation* (1992)

The public schools faced an unsettling predicament as the reign of Ronald Reagan and Margaret Thatcher waned in the late 1980s. Education spending per pupil had continued to climb, rising by over half between 1974 and 1992 in real dollars (Ladd & Hansen, 1999). Yet student achievement had remained flat overall through the 1980s in the United States, brightened only by modest gains among young children from poor families (Berliner & Biddle, 1995).

This lackluster performance by America's students – and presumably their teachers – sparked calls for tighter school accountability or infusing the sector with market competition. Reagan pushed unsuccessfully for federal tax credits that would have subsidized parents already sending their children to private or religious schools. Three states created publicly funded vouchers largely aiding low-income families who opt for Catholic schools. And the fledgling charter school movement, first lent credibility and public dollars by legislatures in Minnesota and California, spread across the land in the 1990s (Fuller, 2000).

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
**Research in Sociology of Education, Volume 16, 1–29**

**Copyright © 2008 by Emerald Group Publishing Limited**

**All rights of reproduction in any form reserved**

**ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16001-3**

Other Western nations, especially the United Kingdom, were a decade ahead of Reagan's conservative curve. Disaffected by its 1970s experiment with progressive education, Thatcher's education advisors imported a fresh corporate model of organizational reform, similar to how nineteenth-century America's progressives had borrowed the grand size and bureaucratic form of modern schooling (Tyack, 1974).

Public support remained strong for somehow fixing the public schools during the Reagan–Thatcher years. But even moderate political leaders concluded that more funding would not likely spark achievement gains in the absence of stiffer accountability. And conservative leaders did not want the diagnosis to seep out beyond the school's walls, hesitant to take on the corrosive forces of urban decay and family poverty. The political framing of the alleged problem would come to emphasize a lack of productivity – insufficient “performance” of student and teachers alike – all within the school's institutional borders.

## THE POLITICAL IMPERATIVE OF PERFORMANCE

Given this casting of the problem, the logical question by the late 1980s had become, how should government craft policy tools to motivate stronger efforts by local educators? A variety of central governments in the West had tried to lift children's learning curves through new funding for particular categories of students, along with tighter regulation of how these dollars must be spent. But this assumed that legislators and education bureaucrats knew how to best organize instructional “inputs” and social relations inside classrooms. The conceptual breakthrough with the new buzz around *standards-based* or *performance-focused reform* was that government would concentrate on clarifying learning *outcomes*, leaving local educators to tailor school inputs and pedagogical practices. (Several chapters in this volume show how, in fact, central governments have difficulty resisting the exercise of control over output standards *and* input mixes.)

The past generation of standards-based (or top-down) accountability policies has aimed to alter the deep-seated institutional habits that characterize the inertia of schools in particular national or state contexts. Public schooling in many societies has reflected a mix of local control and centralized rules, going back to the modern state's initial funding of mass education three centuries ago. Village councils often built and ran modest schools, or central states slowly incorporated church-run schools into a loose, government-financed network (Garnier, Hage, & Fuller, 1989). But

the notion of a unified school *system* did not emerge in the United States until administrative progressives in the late nineteenth century began to consolidate school districts into large units, encompassing many “school plants” into more efficiently run networks. Still, state education departments remained small. Municipal elites and professional managers – not central government – were to advance economies of scale, sort children into ability levels, and standardize administrative routines.

Both the GI Bill and the civil rights movement bolstered faith in central government’s capacity to widen access and equalize school quality. The creation of so-called categorical aid programs was fed by the succession of civil rights movements to aid blacks, Latino, and then disabled children from the mid-1960s forward. The massive Title I program, embedded in the federal education act came to serve millions of students from low-income families, along with federal support for bilingual and special education initiatives. Title I would eventually become the major carrot waved by Washington to press for local compliance to the *No Child Left Behind Act of 2001* (NCLB). This widening river of categorical aid would flow through state education departments, each program replete within regulations, compliance monitoring inside schools, and a vocal set of constituencies. California’s state education department currently oversees more than 130 categorical aid programs, sending dollars and rules down to the state’s local school districts, comprising over one-third of all school spending (PACE, 2006).

Centralizing tendencies within states were further hurried along as civil rights groups successfully pushed to equalize spending across districts, be they rich or poor. Influential court action in California, New Jersey, and Texas – mainly in the 1970s, a generation before the accountability press – aimed to equalize inputs among districts, presumably advancing parity in student achievement. Greater equity in tax burdens felt by poor versus affluent families required decoupling school finance from local property taxes. This would come to shift control over tax revenues from local boards to governors and state legislatures, setting the stage for more intense centralization under accountability regimes. The emerging bargain was that new spending would become available as state government doled out tough love for local educators.

Earlier investigations showing that fresh dollops of spending and inputs alone would not likely equalize children’s learning were largely ignored by the 1980s (Coleman et al., 1966; Rutter, Maughan, Mortimore, & Ouston, 1982). We would later discover that spending equity among districts only modestly equalized resources allocated among schools within urban and

ageing suburban districts. But as New Democrats, like southerners Bill Clinton and Richard Reilly (the South Carolina governor who would become Clinton's education secretary), realized the political advantage of getting tough on teachers and urging strong productivity, both input strategies and equity concerns waned somewhat in Washington.

New dollars for schools would come under Clinton and Bush administrations alike, much of it targeted on lower-income neighborhoods. But the dominant frame alleged that stronger performance could be squeezed from the public schools within existing budgets, and that educators and students should be punished if they failed to respond to more demanding standards and an expanding maze of performance rules. It was no longer legitimate – in polite centrist political circles – to suggest that family poverty or institutionalized disparities among communities were the underlying problem. Nor could local school boards, left with little control over revenues and presumably captured by education interests, be trusted to exercise local accountability.

### **CENTRALIZING RULES TO SPARK LOCAL MOTIVATION?**

Top management writers like Tom Peters and Robert Waterman (1982) offered a timely gift to moderate politicians on both sides of the Atlantic – especially those who aimed to “reinvent” government, from Al Gore to Tony Blair. Management theorists were intrigued with companies that, while quite profitable, had eschewed the old industrial model of routinized production technologies and top-down controls implemented by a thick layer of middle managers. Firms like Hewlett-Packard and Johnson and Johnson set forth crisp production goals for local managers and task groups, according to the management gurus. Then, corporate headquarters granted local production units wide latitude to devise better ways of meeting output goals, often encouraging local inventiveness. “Organizations that live by the loose-tight principle are on the one hand rigidly controlled, yet at the same time allowed autonomy, entrepreneurship, and innovation from the rank and file” (Peters & Waterman, 1982, p. 318).

Public sector reformers were attracted to this simple notion of “loose-tight” organizing, as they pitched for more efficacious action by government with less bureaucratic control and micro-management of local reform efforts. Bill Clinton, who chaired the bipartisan National Governors



Association in the 1980s, liked to cite Osborne and Gaebler's (1992) book, *Reinventing Government*, a volume that "should be read by every elected official ... this book gives us the blueprint," Clinton said. The restructuring of government came to be fused with Clinton's centrist policy agenda, stressing that "expansion of opportunity, not bureaucracy, (that is) reinventing government, away from top-down bureaucracy of the industrial era to a leaner, more flexible, more innovative model appropriate for the modern, global economy" (Clinton, 2004, p. 361).

Clinton and fellow moderates had not given up on boosting school funding, nor quietly pushing for greater equity. But they aimed to blend the old Left's faith in central regulation with the pitch by business moderates who agreed with the Right that local schools suffered from hazy aims and suffocating bureaucracy. The emerging top-down accountability strategy – manifest in crisp learning standards enforced through state or federal monitoring – banked on a new bargain. Voters would presumably back more robust spending if they saw stronger returns on their current investment.

State governors and legislators would take the lead in advancing accountability reforms through the 1990s. The unprecedented education summit called by President George H. W. Bush in 1989, held in Charlottesville, Virginia under the gaze of Thomas Jefferson, was attended by nearly every state governor and co-chaired by Clinton. The gathering would ratify an unprecedented set of national goals for education, while leaving with the states the responsibility for crafting stiffer accountability measures, systems for tracking student performance, and perhaps rekindled finance reforms.

Veteran congressional staffer Christopher Cross (2004) attended subsequent negotiations over how to implement the Charlottesville accords, held at the Bush White House. Focusing on one ambitious goal – that American students would be first in the world in math and science achievement by 2000 – Clinton arrived with a list of funding strategies and suggested legislation, focused on input policies. Bush's point person, Roger Porter, preferred in contrast to design a better way to track student performance over time, and Washington would simply monitor states' progress in meeting what became seven national goals for education.

The real action on output-focused reforms was well underway out in the states, as federalist Jefferson would have enjoyed observing. Forty-three states had taken steps by 1993 to transparently hold schools accountable for the performance of their students. This included tightening accreditation practices, delineating clearer learning standards, expanding state testing

programs, and doling out rewards or sanctions to schools based on performance (Elmore, Abelman, & Fuhrman, 1996). Many states, redoubling efforts to address equity, were focusing additional resources on the so-called “low-performing” schools, whether the prior background of children and families was taken into account or not.

Twenty-eight states were compiling test score data in publicly accessible ways by the end of the century, focusing attention on the performance of individual schools, not only gauging the statewide progress of children (Goertz & Duffy, 2001). Several southern states were experimenting with monetary rewards for schools or teachers who appeared to be raising student performance, getting the incentives right as economists had urged (Clotfelter & Ladd, 1996). Kentucky had discarded norm-referenced testing, where the state’s children were compared to national norms, opting instead to define the share of students who performed at “proficient” levels in reading, math, and other subjects. South Carolina, as far back as 1984, had begun to think about shaking-up the entire school system – indeed, the phrase “systemic reform” was creeping into the parlance – by raising teacher pay while boosting course requirements, instituting a new testing scheme, and funding performance incentives for principals who succeeded in boosting achievement.

## **TAKING STOCK: THE DILEMMAS AND LOCAL EFFECTS OF ACCOUNTABILITY**

A generation has passed since the American states began pressing centralized accountability policies, complete with distinct learning objectives, stiffer performance standards, and centralized monitoring of progress. A similar model of accountability, focusing more on school outputs and less on equitable investment, has swept across the United Kingdom, other parts of Europe, and many developing nations where more localized forms of schooling had earlier dominated the institutional landscape. Centralized curricula and high-stakes testing, of course, have long characterized educational institutions in much of Asia and some European countries, from China to India to Germany (Lockheed, 1997; Fuller & Rasiah, 2005). Thus the steady centralization of school control – from what is to be taught, with what resources and through what legitimate pedagogy – is not always a new dynamic when looking cross-nationally.

What is intriguing about the U.S. case is that top-down accountability policies attempt to reshape what had been a far-flung decentralized governance structure. And the central state – even when controlled by left-of-center moderate political leaders – continues to battle powerful education interests to push for stronger motivation and performance inside local schools. Odd bedfellows abound in the current debate in Washington over how to recraft and reauthorize NCLB. Civil rights groups, moderate business lobbies, and President Bush remained arm-in-arm through 2008, pushing the Congress to protect the core principles and myriad rules contained within the “No Child” law. Opposed to merely tinkering with the law are the national teacher unions, the National Governor’s Association, and local school boards (Smith & Fuller, 2007). The essential debate over how to improve schools has moved up to the national government, a central forum that represented an illegitimate stage for such contestation even a decade ago.

Many educators on the Left pitch the virtues of liberal learning for civic participation, a fundamental tenet pushed by Jefferson, John Dewey, and early democrats. Yet equity proponents, rooted in the history of school finance reform and federal intervention to aid poor families, seek ever more muscular centralization. The Right continues to press for greater productivity and less bureaucracy, at times advocating for wider institutional diversity, as with charter schools or vouchers for religious schools. Amidst this swirl of cross-currents, the basic contradiction remains stark: political leaders of many stripes and colors continue to press for top-down accountability measures while backing the core notion of liberal learning. Somehow the next generation will learn to think for themselves, to engage the civic sphere, to act as intelligent market actors. But they will acquire this knowledge and necessary social inclinations via sacred texts and uniform pedagogical practices handed down from the central state.

This volume aims to take stock of the empirical benefits and such dilemmas stemming from top-down accountability – along with diminishing achievement gains – that have come to light over the past two decades. Government *has* reinvented itself when it comes to who defines what all children must learn, who shapes the work of teachers, and the materials they must employ. But what are the sustained effects of strong state accountability measures inside schools? How do teachers interpret, and adapt to, these performance-obsessed policies? And does this politically advantageous strategy yield the intended effects of boosting the motivation of teachers and the learning curves of students? These are the fundamental questions addressed by our contributors.

Given this historical backdrop, I next turn to a simple framework for analyzing the organizational mechanisms and local effects of top-down accountability. This analytic strategy has three parts. First, we review core elements of the *original accountability model*, as crafted by its early architects. The historical conditions sketched above served to make standards-based reforms both politically appealing for the central state and educationally sensible (at least in their original form). I also highlight the specific organizational mechanisms – or postulates of the first accountability designers – that were to kick-in and energize stronger performance by teachers and students alike. This frame emphasizes how implementation of top-down accountability, first by the states and then by Washington, was as much a political compact (tough love in exchange for new dollars) as it was a sound strategy for innovation or motivational lift inside schools. A variety of state actors aim to be seen as decisive, tough-headed reformers – whether their organizational change strategy is ever felt on the ground or not.

Second, I review evidence on the medium-term *benefits* of government's redoubled focus on achievement. National assessment scores did climb in states that pursued more hawkish accountability policies during much of the 1990s, at least within elementary and less so in middle schools. This progress was remarkable in math and at times significant with regard to reading proficiency. Gaps between children of poor and middle-class families also narrowed during the 1990s. Yet sustaining this momentum has proven to be difficult. Achievement gains seen in the 1990s flattened-out after the 2002 enactment of NCLB. Our contributors tickle the question of whether performance gains were damped by stricter federal intervention, or whether federal involvement simply came on too late in this policy cycle – after the benefits of state-led accountability had played out.

Third, we take stock of the extent to which top-down accountability may continue to alter *social* relations and *norms about knowledge* inside classrooms. This volume's contributors advance our empirical understanding of this crucial question, one marked more by ideology and political philosophy than hard evidence. They also empirically speak to the point that intermediary organizations – operating between the state and school – play a pivotal role, interpreting and mobilizing the tools available under new learning standards and incentives to raise student performance. This intermediary role involves school district leaders and technicians and non-governmental organizations (NGOs), as they help teachers and principals make sense of this brave new world of state involvement, its hazards and possibilities.

*Recrafting the Institution: Power, Capacity, and Pedagogy*

The control of schools by village councils or city mayors had long been seen as a virtue, not a problem, across a variety of Western societies well into the modern era. In fact, liberal forms of human development and individual rationality required democratic governance at the grassroots. Faith in local control did not necessarily ensure a rainbow of diverse forms of schooling, before or after Horace Mann pressed for “common schools” in the second-third of the nineteenth century. What a one-room school house or paid tutor was to look like had become institutionalized. Nor did decentered control yield equitable spending, since school revenues were derived from each community’s property values.

Still, the merits of local control did not come under stiff scrutiny until the segregationist South stung America’s dormant civic consciousness in the late 1950s. Then, the steady rise of Asian economies and sudden oil scarcity in the 1970s, manipulated by a mysterious group calling itself OPEC, revealed an American economy that was no longer preeminent and even quite vulnerable. The very metaphor of a nation “at risk” of decline on a grand scale would become the title of an influential federal report on education, put out in President Reagan’s first term. America’s largely decentralized and apparently stagnant non-system of public education was partly to blame for the country’s perilous economic situation. Bill Clinton, through the 1990s, would echo the argument that stronger public schools would boost the economy and advance upward mobility. Economic or anti-poverty policies to aid children’s development were no longer a priority. The causal arrow went the reverse direction. The school institution must first be fixed, tightened-up; this would then drive economic growth according to the self-proclaimed New Democrats.

The post-war spread of higher education had propelled millions of black graduates, and a rising number of Latino and Asian youths, into the middle class. Yet forced bussing in many cities had spurred white flight, leaving behind few middle-class families and a collapsing tax base. The flight from dismal city schools would feed conservative support for tuition tax credits, publicly funded vouchers, and, by the early 1990s, charter schools. And much of the political energy behind the latter two movements came not simply from conservative elites, but also from lower-income parents who desperately sought new options for their children (Fuller, 2000). Britain had long supported grant-in-aid schools, traditionally supporting Anglican schools, similar to public support of diverse, even sectarian schools by Scandinavian governments. So, the Right argued,

governments should encourage school choice and nurture market competition, not simply spend more on a monolithic model of schooling that displayed little capacity to improve.

### *The End of Loosely Coupled Schools?*

Meanwhile, students of school organizations, including Stanford sociologist John W. Meyer and his intellectual disciples, had detailed how layers of the school institution were *loosely coupled*, allowing principals to buffer outside pressures from parents and district officials alike, enabling teachers – good or bad – to continue their classroom practices behind closed doors (Meyer & Rowan, 1978; Gamoran, Secada, & Marrett, 2000). Principals and teachers often engaged in symbolic compliance to interventions from the outside, a strategy made possible by the rituals and routines that signaled to outside constituencies that real schooling was occurring, and by the lack of transparent performance data. The notion that highly institutionalized organizations like schools could draw more and more public funds while decoupling from performance demands infuriated conservative critics. The loose-coupling metaphor would also influence early accountability architects, some who shared the intellectual roots set down by the Stanford line of analysis.

John Chubb and Terry Moe (1990), describing the post-1960s politics of many school districts, shed light on surrounding many local school boards and principals, arguing that they had become penned-in by the rules governing categorical aid, teacher unions who elected their own members to school boards, and strong parent constituencies, such as those of disabled children or those fighting for bilingual programs. Local districts had come to be self-serving, according to Chubb and Moe, protecting bad teachers, pressing a senseless array of fragmented programs, and failing to award principals the discretion necessary to improve their schools. Achievement remained stagnant for years in many locales, but no one held the authority or chutzpah to press for stiffer accountability. Thawing-out the hardened and often ineffective practices of teachers and principals would be challenging, according to Meyer, since the technology of teaching is so inexact and carried out within the confines of isolated classrooms. But for Chubb and Moe, the issue was simply one of power. And if parents were free to operate in an educational marketplace, the power of education interest groups could be restrained.

The early architects of standards-based accountability, writing in the late 1980s in the wake of these intellectual currents, defined the public problem in appealing fashion. And the resulting remedies offered a fresh, hopeful role for the central state, while no longer arguing that money alone was the answer. The newly constructed problem rested with local educators themselves and the incoherence of their local organization. In the absence of “systemic reform” the loosely coupled layers of the school institution and increasingly fragmented programs would protect lousy teachers and ineffectual principals. Local school boards could not exercise strong accountability. The entire system – from specifying statewide learning objectives and testing programs, to dispassionately monitoring of progress – needed fixing. Governors and state legislators would override opposition from education interest groups, convincing them that little new money would become available until performance improved.

But how could the federal or state governments revamp the institution without expanding bureaucracy at the top of this newly rationalized *system* of public schooling? Government itself was to be “reinvented” in a leaner and (somewhat) meaner form, creating a tidier division of labor between central and local organizations. Michael Cohen, schooled in organizational behavior at Stanford University, had become a policy advisor at the National Governors Association, the bipartisan group that Clinton co-chaired in the 1980s. Cohen (1990), along with Stanford dean Marshall Smith (Smith & O’Day, 1991), sketched the elegant corner stones of standards-based accountability policies. Smith would become Clinton’s deputy secretary of education; these were not the idle doodlings of idealistic academics. Nor did core elements of state-led accountability – under their original blueprint – reflect a top-down theory of how to motivate change down under.

These policy designers did borrow the European notion that the American states must delineate crisp curriculum standards, grade by grade, when it came to the knowledge that all children were to master. This was a key element of Thatcher’s own version of tight-loose accountability, pressed in the early 1980s. But in the U.S. context it was the states – not the federal government, according to Cohen, Smith, and O’Day – who should define the learning goals and track student progress. And local educators were to be left to their own devices in mixing resources and shaping pedagogical practices to raise achievement.

Their tight-loose organizational strategy departed from the post-1960s emphasis on aggressive state regulation of proliferating categorical aid programs. Reinvented government agencies would focus mostly on setting

output expectations rather than engineering inputs. Cohen also urged states to emphasize complex learning skills, rather than being satisfied with the transmission of knowledge in basic subjects, such as reading and math. He foresaw government's tendency to routinize knowledge and forms of learning easily gauged by standardized testing. Unlike Thatcher's more centralized approach, mimicking the likes of Germany and Japan, the original American version of accountability was quite federalist, placing authority and output-related tools in the hands of governors and state school chiefs, not under the control of Washington regulators.

### *Delicate Theory Meets the Bureaucratic State*

What these early architects failed to predict was the downstream inclination of many states to reject more complex forms of learning, mandate particular curriculum packages, and encourage didactic instruction (e.g., Stecher, 2002). Remember that the original rendition of innovative corporate organization, advanced by Peters and Waterman emphasized central management's laser-beam focus on setting performance standards and tracking progress, while leaving the *technical core* of production under the control of local managers and skilled workers. After all, the new corporate structure aimed to motivate stronger effectiveness and high quality through innovation – largely by raising the capacity of local managers to combine inputs and motivate staff with less-routinized control from above.

Cohen and colleagues endorsed this conception, stressing that challenging and complex forms of instruction would be advanced through demanding curricular standards. Drawing on new developments in cognitive science Cohen (1990, p. 261) argued that “skilled readers, even at the elementary grades, are able to comprehend what they read not simply because they have acquired the basics, but because they intuitively and automatically rely on what we think of as higher-order skills.”

This realignment in the division of labor was to encourage local educators to advance deeper comprehension and student practice with concepts and ideas, including stronger analytic skills. Local districts would be awarded resources to advance the instructional capacities of principals and teachers, essential in building capacity to address the more ambitious curricular standards. The state would articulate performance standards and expectations for achievement growth, just like an avant



garde high-tech firm, then improve the working conditions and skills of the local artisans – teachers and principals – to inventively manage inputs and pedagogical innovation.

Fellow reformers talked of a more even and energizing division of power between central levels of the state and local education authorities. This was not to be a top-down, Fordist conception of command and control. Power was not seen as a zero-sum game (Fuhrman & Elmore, 1990). Different levels of the institution would each focus on how to meet the learning standards, how to build a more professional teaching force. In retrospect, this pillar of performance-focused reform feels a bit naïve, as states and then Washington pushed to set learning aims *and* dictate instructional methods through routinized curricular packages. The state’s habitual focus on standardization and efficiency was applied to student testing – exams that could be scored more cheaply if they asked students to report on bits and pieces of known information. In turn, textbooks were aligned with new tests, reinforcing the bit-by-bit drilling of knowledge into students’ heads. But we are getting ahead of the story: several of our contributors detail how top-down accountability polices have diverged from the original, more elegant policy model.

Another key feature of the original policy replica stressed that student testing should yield transparent results, revealing the effectiveness of local schools. Many states moved to administer standardized tests at each grade level, beginning in the second or third grade. Some moved to a mastery conception of student assessment, like the Carolinas and Texas, reporting on the percentage of students who displayed basic, proficient, or advanced skills in basic subjects. Testing would involve high stakes for students as well as local educators, including new exits exams that high school pupils must pass before graduating.

Methods for tracking change in aggregate student performance – which have come to be equated with the school’s performance – were devised by states as well. Some then tied the tracking of growth to cash incentives for schools or teachers who appeared to boost achievement, although the data requirements necessary in building a true “value-added” gauge of a school’s discrete effects after factoring in prior achievement and student mobility remain daunting. State and federal policy makers have proven far more willing to dole out sanctions to low-performing schools than to reward schools that display growth. It is much cheaper. Again, the original architects were hopeful that states would allocate positive incentives to effective schools and teachers, a policy plank that also proved naïve when state budgets went south.

*A Federalist Model of Accountability*

The original architects, like Cohen and Smith, expressed the Hamiltonian view that local authorities must be held accountable by central levels of government as one way to raise organizational effects. But they also emphasized how the model could be situated in a federalist structure of politics: governors and state school chiefs would lead the charge in corralling their disparate networks of local schools into more integrated statewide systems. The modern-day American reformers invoked Jeffersonian ideals as well, or perhaps John Locke's conception of a thin yet rational state: promising that centralizing regulation at the state level would ensure demanding yet tailored forms of curriculum, sensitivity to local preferences, and capacity for organizational innovation. After all, it was the states that now raised the bulk of necessary tax revenues for the schools, not local school boards or the federal government.

This volume's contributors speak to an overarching question: What happens when an elegant theory of organizational change is taken up by state actors under pressure to raise the effectiveness of local institutions. Yet the state often remains constrained by the regulatory and bureaucratic habits that have characterized its apparatus for over two centuries. Indeed, it is the ghost of Max Weber and his scholarly autopsy of top-down bureaucracy, reminiscent of the actual accountability regimes, that have come down from state governments and now from Washington under NCLB. Central levels of government in the immediate past have proceeded to mandate certain curriculum packages, buy textbooks that mirror atomistic notions of knowledge, and sanction "failing schools" based on cross-sectional snapshots rather than reliable data on achievement change over time.

**TEACHERS AND PRINCIPALS RESPOND:  
ORGANIZATION OF THE BOOK**

Despite all the policy talk and palpable centralization of school rules, evidence remains spotty on how local educators respond to accountability incursions. This volume's contributors illuminate the daily lives of teachers and principals on the frontlines, interpreting and reacting in concert with colleagues, often with limited resources and thin organizational support. Each study attempts to piece apart the policy mechanisms inherent in

accountability policies – from better achievement information about student learning to high-stakes exit exams – to dig into observable effects on teacher motivation, pedagogical change, and pupil performance.

*Liberal Learning in Centralized States*

Our contributors also examine how political leaders and local educators are finding themselves on the sharp horns of a practical dilemma – one that is far from new within democratic societies, recurrently worried about their economic and political position in a competitive world economy. In the West, many activists accent the postulate that the sacred individual must acquire the skills and moral commitments necessary for bolstering self-reliance, the capacity to achieve in the economy and community as an independent creature. But over the past generation, as these ideals of liberal learning continue to be voiced, social relations and learning inside our schools are increasingly shaped by central dictates: what all children are to learn, in what language, and largely through didactic transmission of known facts.

The debate over this wrenching contradiction will continue and likely intensify across Western societies. Empirical research can play a crucial role, informing this contradiction. Beyond describing how top-down accountability has taken shape in the states and from Washington, lines of empirical work are painting a more detailed picture of whether this policy strategy is yielding positive effects and the intended causal mechanisms that are breaking down, from uneven school capacity and will to focus on students left behind to insufficient resources. These lines of empirical work – advanced further by our contributors – move us beyond the ideology contention around centralizing school policies to understand the conditions under which top-down accountability works, and when it does not, where does implementation fail?

The Western state's recurring drift toward Weberian-like regulation of local behavior – even how we thinking about human learning and child development – has long worried social observers, going back to Jefferson (Madison, 1961) and Locke (1964) Central states have historically aimed to incorporate, even civilize groups on the periphery of civil society, be they recent immigrants, those who speak a different language, or those who signal low social status. Secular schooling is the main agent of this political project of acculturation. Especially in pluralistic societies, the school offers a shared public space in which all children allegedly win the opportunity to

acquire requisite beliefs, job related skills, and democratic understandings which yield rights and comforts linked to membership in the nation-state. Individual capacities for civic engagement and informed market action have long been prized in democratic-capitalist societies.

Yet there is the deepening contradiction with contemporary schooling, the institution that is to impart the civic republican ideals of independence and social citizenship, in Linda McClain's (2006, p. 88) words. Coming off the metaphor of "a nation at risk" and the attraction of novel role for the state, accountability hawks now focus on criticizing, then renewing the public schools. These political moderates seized more power at central levels of government from Bill Clinton to George W. Bush, each with the best of intentions. An activist state was to close the achievement gap and demonstrate that a smaller, shrewder government could work. But are centralizing states more capable of nurturing individuals' capacity to engage in "self-government," or to build from the colorful social foundations found across diverse local communities? To put a sharper point on one empirical question: Do top-down accountability reforms advance students' (and educators') democratic inclinations to engage social institutions and the common good?

Some argue that the cause of anti-poverty and social justice should trump worries over the Western state's centralizing drift. Long-term public problems, from global warming to economic sustainability, will certainly demand strong action by central circles of policy makers. But it does not necessarily follow that local institutions, from schools to housing development, should be subjected to dictatorial controls – especially when inculcating a spirit of civic engagement is so essential to energizing robust institutions.

Active-state advocates have long argued that only central government holds the political and fiscal power to combat the structured inequalities that beset most capitalist societies. Progressive disciples of Hamilton, including Keynes and Franklin Roosevelt, argued that only the central state held the authority to hold elite interests accountable and redistribute wealth. And the press for greater equity requires that government put in place universal gauges of progress, like the barometers afforded by standardized tests. This is one reason why key members of the civil rights community and school finance litigators support more standardized assessment of student learning. But will the didactic drilling of known facts truly advance youths' capacity to think for themselves, to analyze situations, and to examine alternative courses of action? We are back to the paradox of how many political leaders hope to aim to advance liberal learning via centralizing policy tools.

*Fading Empirical Effects from Top-Down Accountability?*

Initial results stemming from state-led accountability efforts – often interwoven with school finance reforms – were quite encouraging. Achievement in math and reading climbed in the fourth and eighth grades during the 1990s, and racial gaps narrowed dramatically (about one grade level). Fig. 1 shows patterns at grade four. Achievement gains were indeed steeper for students in states that pursued more aggressive accountability reforms, according to national assessment results (Carnoy and Loeb, 2002; Hanushek & Raymond, 2005).

But reading curves rose up onto a dusty plateau by 2001, and math achievement has grown at a slower rate since passage of NCLB in 2002 in grades four and eight. The performance of high school students has remained largely immune to accountability reforms, with the exception of more students taking algebra and advanced math courses. America’s students did show distinct gains in reading and math when sitting for the 2007 National Assessment of Educational Progress (NAEP). Yet this was the first consistent shift upward since the 2002 enactment of NCLB, and racial gaps have not discernibly narrowed (Lee & Wong, 2004; Fuller, Gesicki, Kang, & Wright, 2007). It remains unclear whether the highly centralized NCLB regime dampened earlier progress made under state-led reforms, or whether Washington just entered the accountability game too late in the policy cycle. The accountability press has long shown clearer benefits when it comes to younger children’s basic reading and math skills, as opposed to advancing complex cognitive skills in higher grades. These

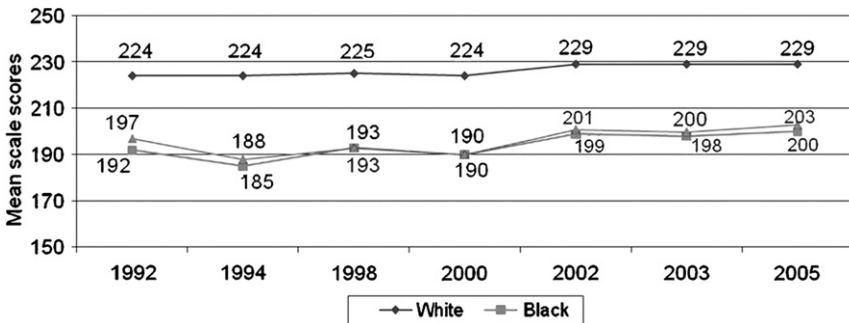


Fig. 1. U.S. National Assessment of Educational Progress (NAEP) Mean Scale Scores by Ethnic Group, 1992–2005.

benefits, once realized, may simply be followed by diminishing marginal returns to static accountability policies.

Some states continue to report gains on their own reading and math tests (Jennings, 2007). Yet debate continues over whether this represents substantive buoyancy in learning or simply more effective teaching to the test. Most states set their own bars for defining “proficiency” much lower than the federal NAEP standards. This allows more students to clear the hurdle when gauged by state tests. Several reports have come to light showing how states manipulate their proficiency bars, adjust item scores in ways that make performance comparisons impossible across years, and test forms not-so-secretly circulate around states, encouraging teachers to sharpen test-prep activities (Stecher, 2002).

Take the case of reading proficiency at the elementary school level. The NAEP governing board has estimated the share of fourth-graders proficient in reading at 31 percent in each of the three years of federal testing which have followed passage of NCLB. Despite this plateau in federal results, our earlier study of state test score trends found that governors and state school chiefs are still reporting annual gains of 1.9 percentage points in the share of students proficient since 2002 when gauged by their own tests (Fuller, Gesicki, Kang, & Wright, 2007). State officials stick by their persisting gains, arguing that NAEP exams are not aligned to any state’s curricular standards and that its proficiency bar is set unrealistically high. But it remains an open question as to why state reading scores, and to a lesser extent math scores, continue to climb with no corresponding bump in national assessment results in many states.

### *Altered States: Tightening Social Relations Inside Classrooms*

One key aim of accountability hawks is to align learning objectives set by states to the actual curriculum implemented by classroom teachers. As state curriculum panels list the pieces of knowledge that all fourth-graders should know in language arts, or eighth-graders should understand when it comes to math, teachers presumably would become more proficient in transmitting these bits of knowledge. At least this was the working hypothesis of skeptical scholars who began studying the effects of accountability policies.

Initial empirical studies reveal a more complicated picture. Linda McNeil and Angela Valenzuela’s (2000) investigation in Texas classrooms offered rich evidence of how many teachers feel that their work is being deskilled, as they are forced to transmit constrained forms of knowledge, often simply

preparing their children for standardized tests. Yet when Elisabeth Woody's team (Woody, Buttles, Kafka, Park, & Russell, 2004) spent extended periods of time inside eight diverse California schools, they found mixed interpretations and classroom responses by teachers. Some believed that accountability measures had usefully focused school staff on problems faced by low-performing kids; test results broken down to specific domains and skills provided clearer diagnostics. Several teachers reported improved collaboration, efforts to ensure that all teachers within a certain grade level were squarely addressing the state's curricular standards.

Borko and Elliott (1999) also report how some teachers do innovate and experiment with fresh pedagogical approaches when earlier practices fail to boost student performance. Mintrop and Trujillo (2007) found stronger achievement gains in schools where teachers reported substantive responses to accountability demands. Our survey of California principals revealed that many schools are lengthening instructional time, including after-school tutoring and Saturday classes to serve low-performing kids (Fuller, Loeb, Arshan, Chen, & Yi, 2007).

Still, negative fall-out is often felt by teachers in the wake of top-down accountability, according to Brian Stecher's (2002) review of statistical evidence and qualitative case studies. Several classroom studies have found that teachers comply by narrowing the range of knowledge and breadth of discussion that many topics receive – if they are domains or activities that contribute little to test scores. One earlier study found that Arizona schools shrunk time spent on social studies, writing exercises, and science (Smith, Edelsky, Draper, Rottenberg, & Cherland, 1991). Similar declines in instructional time dedicated to non-tested subjects or topics have been shown in three other states (Koretz, Barron, Mitchell, & Stecher, 1996; Stecher, Barron, Chun, & Ross, 2000). Focusing on math instruction, Romberg, Zarinia, and Williams (1989) surveyed a nationally representative sample of teachers and found increased emphasis on elemental skills and paper-and-pencil computation, along with diminishing attention to projects and more complex work with calculators. Stecher reviews several studies that detail how teachers – in the face of accountability pressures – spend many days on test preparation, fitting known facts simply to the format of standardized exams.

### *The Next Act: Keep the Faith or Recraft Accountability Policy?*

In sum, we know that assertive top-down accountability policies have spurred stronger achievement, at least in basic skills for elementary-age

children, yet over brief stretches of time. Initial work also shows that discrete policy tools, such as more demanding course requirements or aligning instruction to curricular standards can pay-off, if one's conception of human learning is tied closely to knowledge gauged by standardized tests. At the same time, accountability hawks are feeling the limits of their generation-old strategy, given the flattening of test performance in reading and slowing growth in math scores seen in recent years.

### *Illuminating Action across Organizational Levels*

Still, we remain in the empirical dark when it comes to explaining the tandem mysteries of strong gains in the 1990s and, now, the dry plateau onto which student performance has climbed. This is more than a single black box; it is a stack of blocks in which obscured organizational mechanisms may operate at differing levels of the education institution (Rowan, 2006).

The chapters that follow focus largely on life down under, inside schools, seeking to understand how teachers and principals are reasoning about, and adapting to, top-down accountability policies. Our contributors push to understand how accountability regimes – an enigmatic blend of federal NCLB and parallel state policies – are being felt by educators inside schools, often mediated by district leaders. These scholars paint colorful and detailed portraits of evolving features of life inside schools and classrooms, drawing from interviews, surveys, and qualitative inquiry. The chapters collectively help to pull apart the moving parts of top-down accountability policies, from looking at the effects of high school exit exams to being deemed a “failing school.”

These authors are sensitive to the modest magnitude with which state or district-level actions yield change in teachers' daily motivations or pedagogical agility. The uneven support or regulation advanced by district officials, discovered by some research teams, offers one explanation for why top-down standards and controls are yielding uneven returns.

Taken as a collection of fresh empirical findings, the chapters inform key theoretical and practical questions related to how the school institution's habits persist and which discernible shifts in organizational behavior may stick over time.

- *Which policy tools survive as they move through organizational levels and actors?* The early architects of standards-based accountability believed that the strategy would provide crisp learning aims and closely aligned instructional materials. Clearer data about student performance would



then provide a tool for teachers to address gaps. But are these tools understood and used by local educators in the ways that the policy theorists intended? Accountability proponents assume that novel agency stems from central government, yet several contributors show how district leaders may offer original action when it comes to blending rules and resources.

- *Which levels of the school institution gain legitimacy, resources, and organizational mechanisms that local educators find helpful?* Top-down accountability was sold politically as a strategy that would bring greater public support and dollars to teachers and their schools. Discrete elements of top-down accountability would make policy makers look decisive and yield new tools for local educators. But we have known very little about the specific mechanisms – curricular structures, teaching materials, training opportunities, or collaborative planning inside schools – that teachers directly experience and find motivating. Critics worry that accountability initiatives have drifted toward prescribed curricula, constraining pedagogical practices, and negative sanctions. But is this what teachers are experiencing?
- *Has the underlying social architecture of the school institution and the classroom discernibly changed?* Early findings during the accountability era consistently suggest that something is changing inside classrooms, whether it is the narrowing of the curriculum or perhaps the recruitment of stronger teachers. But if the magnitude of these effects is small, will they last? Are these improvements sufficiently strong to narrow achievement gaps, or do accountability reforms reproduce stratified patterns? And – moving beyond the technical task of discerning functional effects – do accountability policies push out innovative forms of teaching and learning?

Our contributors also scratch the empirical surface of deeper questions. Within the classroom, are teachers losing control over their craft, showing less inventiveness in how they motivate students? Or, do accountability mechanisms advance professional collaboration and pedagogical gains?

### *Seven New Studies: How Educators Adapt to Accountability Demands*

An intriguing study led by Laura Hamilton leads-off the volume. Her team asks whether teachers' responses to accountability – enacted by individual teachers – are guided by school-level conditions and the leadership of

principals. Hypotheses abound over how between-school differences, such as the class attributes of families or the unequal distribution of quality teachers, may swamp the capacity of individual teachers to improve their practices. Others posit that accountability policies award principals the necessary authority and tools to harmonize teaching practices.

Hamilton's team tested such claims by surveying local educators situated in California, Georgia, and Pennsylvania over a three-year period, 2003–2006. They find that teachers do report distinct adjustments to their pedagogical practices and work with fellow teachers, as they adapt to accountability pressures. But school-level factors, including their respective principal's leadership attempts, explain very little of the variance among teachers' individual responses. Even math teachers (on whom the chapter focuses) are quite eclectic in how they respond to state and federal accountability pressures. The extent to which principals report feeling pressure to raise test scores does help to explain the magnitude with which they shift attention to raising student test scores. But overall, Hamilton's findings suggest that social relations and technical work inside the school remain loosely coupled, despite government's aggressive push for tight "alignment" of state curricular standards and local pedagogical practices.

Thomas Luschei and Gayle Christensen help to explain why the influence of principals remains so uneven. Principals report receiving highly variable guidance or support from their district leaders, despite the diffuse pressure to elevate student performance. Even with a small sample of districts, Luschei and Christensen dug deep via interviews and time inside schools, finding that district staff rarely helped principals to build internal accountability procedures. The use of student performance data remained rare, despite all the additional testing. Distinct-facilitated change inside high schools was especially difficult to find. Neither district leaders nor principals were able to corral disparate department heads or convince teachers to pull in a common direction.

Kristen Gordon's findings yield more upbeat inferences regarding how teachers respond to a pair of policy tools: the adequate yearly progress (AYP) indicators embedded in NCLB, and the program improvement status into which "low-performing schools" are increasingly placed. Based on survey evidence from teachers in Georgia, she found that schools falling short display increasing levels of collaboration. The threat of NCLB sanctions spurred teachers to fight back collectively. One strength of Gordon's design is that she empirically relates the tandem policy tools – the AYP and program improvement mechanisms – to teachers' motivational

levels, and how policy jolts can alter shared behavioral responses inside schools, including meaningful collaboration.

Encouraging results also stem from the chapter by Melissa Henne and Heeju Jang. They inventively track achievement levels for white and Latino elementary students within integrated California schools. Henne and Jang then back-up to ask how reported teacher and principal actions – in adapting to accountability pressures – account for mean achievement levels and the magnitude of ethnic gaps among 111 sampled schools. The conventional school-effects literature has long focused on average student performance among schools. But contemporary accountability policies aim to narrow disparities as well.

They find that Latino students perform at higher levels in schools where teachers report more consistent pressure – with instructional resources and caring supports – received from their principal and district staff. These authors sampled California schools within a narrow range of social-class characteristics, including parents' education levels and home language. They find that Latino students within these schools respond more consistently to accountability pressures than white students, especially when their teachers report support and resources from their principal.

The volume's final three chapters focus on specific policy tools and how benefits stem in part from localized conditions and leadership. Jennifer Jellison Holme reports qualitative findings on how California high schools responded to a newly enforced exit exam. This policy device moved fresh attention and resources within schools to low-performing students. But Jellison Holme uncovered few structural or pedagogical changes in how high school principals were addressing these flagging students over time. Instead, ad hoc tutoring and information sessions were more commonly pasted together at the school's organizational periphery. Again, we see little evidence of crisp organizational change which is more likely to persist over time.

Soung Bae, after discovering a California school district that effectively narrowed achievement gaps among ethnic groups, burrowed deep to understand the organizational mechanisms explaining this success. Much of the answer, she found, rested with district leaders' sustained commitment to inservice teacher development, tied to California's curriculum standards and reinforced through sustained follow-up inside classrooms. The district office revamped its own organization to create and stick with long-term pedagogical support inside teachers' own classrooms. This district led with new resources and professional supports, not rules and punitive measures.

Finally, Michèle Schmidt takes us into a pair of vocational high schools suddenly faced with advancing knowledge contained in Pennsylvania's new

curricular standards. Her study offers a fascinating case of what happens when a specialized form of secondary school is now held accountable to universal forms of academic knowledge – even when students and teachers alike entered these schools with radically different expectations. Schmidt reports how these schools’ basic foundations begin to crumble, even as academic and vocational teachers earnestly attempted to implement the state mandated curriculum. Many students were simply not prepared for the new expectations – not surprising, given that these adolescents opted to enter a vocational program. Schmidt’s study is a cautionary tale for would-be reformers who seek to create inventive forms of high schools – during the current era of top-down accountability.

Taken together, these fresh empirical findings are hopeful in terms of the capacity of local educators – district leaders, principals, and teachers – to respond to accountability pressures. Some teachers try to address novel state curriculum standards and explore new pedagogical techniques. Teachers often report benefits from helpful leadership and fresh dollops of instructional resources, coming from principals or district offices.

But as you explore these chapters be attentive to the consistency with which “alignment” is pressed by school principals, and the extent to which teachers are motivated by this kind of pressure. Why do specific policy tools – from the tracking of student performance to high school exit exams – fail to consistently trigger the kinds of instructional resources and local leadership that many teachers value? If the fundamental structure of schooling – when traveling down from the state capital through district layers and finally to schools – remains loosely coupled (as Hamilton’s findings suggest), what does this say about the tidy alignment imagined by policy makers? These empirical questions deserve more investigation.

As such rich evidence begins to accumulate on the past generation of top-down accountability policies, scholars can more confidently inform technical questions around which policy tools have wielded intended effects on the organization of schools and the practices of teachers. This mounting evidence might also inform the broader historical debate around the benefits and risks of centralizing learning within a liberal-democratic society.

#### *After the Faithful Romance with Top-Down Accountability*

These chapters offer empirical glimpses into why single-minded accountability policies advanced by state and federal governments are yielding

diminishing returns. But do these fresh findings suggest how accountability efforts can be discernibly improved or more efficaciously blended with collateral reform strategies?

We are learning about the discrete accountability tools that appear to pay-off, including strong and consistent attention to learning and teachers' innovative capacity, yet inconsistently delivered by district leaders and principals. The steady use of student data surfaces in these findings as an erratically mobilized tool as well. Certain accompanying resources also appear to help predict learning growth, or at least cross-sectional achievement levels, such as supportive inservice training keyed to specific classroom practices and unrelenting follow-up with teachers in their daily environs.

What remains unknown is whether other resource streams tend to interact with strong benefits for teachers and students, from reducing the size of classes, to instructional coaches, to collaborative planning among teachers. This question of how accountability tools interact with resource flows remains on the theoretical and empirical frontier.

Perhaps most intriguing – looking across these new findings put forward by the contributors – is the inconsistency with which even basic pieces of the accountability puzzle are fit together by local educators. Many district leaders remain inattentive to raising achievement, according to the bulk of teachers interviewed across various studies. A sizeable proportion of teachers report that their principal offers weak or no leadership when it comes to raising student performance, apparently preoccupied with daily brushfires and administrative tasks. While teachers do report discrete responses to accountability demands, including adjusted pedagogy and a surgical focus on low-performing kids, teacher adaptations appear to be largely idiosyncratic, not patterned by inspiring local leaders or collaboration with colleagues.

District leaders appear to *selectively couple* certain activities, penetrating occasionally into the beliefs and behaviors of classroom teachers, while remaining loosely coupled with principals and teachers along other functions (Fuller, 2007). This raises important theory-building questions. Do districts with more rationalized, bureaucratic histories engage in tighter coupling when it comes to implementing accountability policies, compared with more professional districts (often situated in affluent communities)? Do particular technologies, such as highly routinized curricular packages, help to tighten couplings between districts and schools? And when new resources are pumped into schools by states, which activities are more tightly coupled and which become selectively decoupled? Might selective-coupling strategies protect and advance teacher motivation, avoiding the alienating effects of hyper-rationalization?

The chapters that follow also prompt the question, What is the outer edge of the potential effects of top-down accountability in the absence of more resourceful strategies? The Henne–Jang paper is especially eye opening in this regard. Even after sampling California schools that serve similarly working-class communities, achievement levels remained more highly related to parents' education levels and the home backgrounds of children than to the efficacy of accountability programs.

Then there is the philosophical – even ethical – quandary as to whether political leaders should keep preaching the virtues of top-down accountability as achievement effects reach a plateau. At times it appears that politicians and educators alike keep pressing for accountability because they simply do not know what broader, more efficacious strategy might be crafted. The political risks associated with pulling back from stiff accountability measures remain great. But unless this bundle of faithful policies can be implemented more consistently and blended with the resourceful tools identified by our contributors the political utility of top-down accountability may continue to fade.

## ACKNOWLEDGMENTS

My work on the organization and efficacy of accountability policies is funded primarily by the Hewlett Foundation. Appreciation is expressed to Marshall Smith and Kristi Kimball for their unwavering support. Empirical research on how teachers interpret and adapt to accountability regimes, directed by Elisabeth Woody and supported by the Noyce Foundation, greatly informed this volume. Heartfelt thanks are due Amy Gerstein and Ann Bowers Noyce. Woody's collaboration with Jennifer O'Day spurred additional thinking and empirical digging. Research assistants made invaluable contributions to this project, including Soung Bae, Katie Gesicki, Haleh Hatami, Melissa Henne, Erin Kang, Jennifer Russell, and Joe Wright. Completion of this project was made possible by Allison Chen, Doug Kearney, and Sitome Mebrahtu who have provided essential aid at my intellectual home over the past decade, Policy Analysis for California Education. Keith Nitta's doctoral thesis contributed to my understanding of the corporate roots of contemporary accountability policies. The quality of this volume stems, as always, from the steady friendship and support of Emily Hannum, series coeditor.

## REFERENCES

- Berliner, D., & Biddle, B. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. New York: Addison-Wesley.
- Borko, H., & Elliott, R. (1999). Hands-on pedagogy versus hands-off accountability. *Phi Delta Kappan*, 80, 394–400.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24, 305–331.
- Chubb, J., & Moe, T. (1990). *Politics, markets and America's schools*. Washington, DC: The Brookings Institution.
- Clinton, W. (2004). *My life*. New York: Vintage.
- Clotfelter, C., & Ladd, H. (1996). Recognizing and rewarding success in public schools. In: H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: Brookings Institution.
- Cohen, M. (1990). Key issues confronting state policymakers. In: R. Elmore (Ed.), *Restructuring schools: The next generation of educational reform*. San Francisco: Jossey-Bass.
- Coleman, J., Campbell, E., Hobson, C., McPartland, F., Mood, A., Weinfeld, F., & York, R. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.
- Cross, C. (2004). *Political education: National policy comes of age*. New York: Teachers College Press.
- Elmore, R., Abelman, C., & Fuhrman, S. (1996). The new accountability in state education reform: From process to performance. In: H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC: Brookings Institution.
- Fuhrman, S., & Elmore, R. (1990). Understanding local control in the wake of state education reform. *Educational Evaluation and Policy Analysis*, 12, 82–96.
- Fuller, B. (2000). *Inside charter schools: The paradox of radical decentralization*. Cambridge, MA: Harvard University Press.
- Fuller, B. (2007). *Standardized childhood: The cultural and political struggle over early education*. Palo Alto: Stanford University Press.
- Fuller, B., Gesicki, K., Kang, E., & Wright, J. (2007). Gauging growth: How to judge No Child Left Behind? *Educational Researcher*, 36, 268–278.
- Fuller, B., Loeb, S., Arshan, N., Chen, A., & Yi, S. (2007). *California principals' resources: Acquisition, deployment, and barriers*. Stanford: Institute for Research on Education Policy and Practice.
- Fuller, B., & Rasiah, A. (2005). Schooling citizens for evolving democracies. In: S. Fuhrman & M. Lazerson (Eds), *Institutions of American democracy: The public schools*. New York: Oxford University Press.
- Gamoran, A., Secada, W., & Marrett, C. (2000). The organizational context of teaching and learning: Changing theoretical perspectives. In: M. Hallinan (Ed.), *Handbook of the sociology of education*. New York: Kluwer Plenum.
- Garnier, M., Hage, J., & Fuller, B. (1989). The strong state, social class, and controlled school expansion in France, 1881–1975. *American Journal of Sociology*, 95, 279–306.
- Goertz, M., & Duffy, M. (2001). *Assessment and accountability systems in 50 states, 1999–2000*. Philadelphia: Consortium for Policy Research in Education.

- Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24, 297–327.
- Jennings, J. (2007). *Answering the question that matters most: Has student achievement increased since No Child Left Behind?* Washington, DC: Center for Education Policy.
- Koretz, D., Barron, S., Mitchell, K., & Stecher, B. (1996). *The perceived effects of the Kentucky Instruction Results Information System*. Santa Monica: RAND.
- Ladd, H., & Hansen, J. (1999). *Making money matter: Financing America's schools*. Washington, DC: National Academies Press.
- Lee, J., & Wong, K. (2004). The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal*, 41, 797–832.
- Locke, J. (1964). Some thoughts concerning education, 1663. Reprinted In: P. Gay (Ed.), *John Locke on education*. New York: Teachers College Press.
- Lockheed, M. (1997). *Primary education in India*. Washington, DC: World Bank.
- Madison, J. (1961). These departments should not be so far separated as to have no constitutional control over each other (no.48). In: C. Rossiter (Ed.), *The federalist papers*. New York: Penguin Books.
- McClain, L. (2006). *The place of families: Fostering capacity, equality, and responsibility*. Cambridge, MA: Harvard University Press.
- McNeil, L., & Valenzuela, A. (2000). *The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric*. Cambridge, MA: Harvard Civil Rights Project.
- Meyer, J., & Rowan, B. (1978). The structure of educational organizations. In: M. Meyer (Ed.), *Environments and organizations*. San Francisco: Jossey-Bass.
- Mintrop, H., & Trujillo, T. (2007). The practical relevance of accountability systems for school improvement. *Educational Evaluation and Policy Analysis*, 29, 319–352.
- Osborne, D., & Gaebler, T. (1992). *Reinventing government*. New York: Addison-Wesley.
- PACE – Policy Analysis for California Education. (2006). *Crucial issues in education, 2006*. Berkeley and Stanford, CA: Graduate School of Education, University of California.
- Peters, T., & Waterman, R. (1982). *In search of excellence: Lessons from America's best-run companies*. New York: Harper & Row.
- Rodriguez, R. (1992). *Days of obligation: An argument with my Mexican father*. New York: Penguin Books.
- Romberg, T., Zarinia, E., & Williams, S. (1989). *The influence of mandated testing on mathematics instruction: Grade 8 teachers' perceptions*. Madison: National Center for Research in Mathematical Science Education.
- Rowan, B. (2006). The new institutionalism and the study of education organizations: Changing ideas for changing times. In: H. Meyer & B. Rowan (Eds), *The new institutionalism in education*. Albany: State University of New York Press.
- Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1982). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.
- Smith, M., & Fuller, B. (2007). Political leaders must shore up law's shaky foundation. *San Francisco Chronicle*, October 14, p. E1.
- Smith, M., & O'Day, J. (1991). Systemic school reform. In: S. Fuhrman & B. Malen (Eds), *The politics of curriculum and testing: The 1990 yearbook of the politics of education association*. New York: Falmer Press.



- Smith, M. L., Edelsky, K., Draper, K., Rottenberg, C., & Cherland, M. (1991). *The role of testing in elementary schools*. Los Angeles: UCLA Center for Research on Evaluation, Standards, and Student Testing.
- Stecher, B. (2002). Consequences of large-scale, high-stakes testing on school and classroom practice. In: L. Hamilton, B. Stecher & S. Klein (Eds), *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.
- Stecher, B., Barron, S., Chun, T., & Ross, K. (2000). *The effects of the Washington state education reform on schools and classrooms*. Los Angeles: UCLA Center for Research on Evaluation, Standards, and Student Testing.
- Tyack, D. (1974). *The one best system: A history of American urban education*. Cambridge, MA: Harvard University Press.
- Woody, E., Buttles, M., Kafka, J., Park, S., & Russell, J. (2004). *Voices from the field: Educators respond to accountability*. Berkeley and Stanford, CA: Policy Analysis for California Education.

## CHAPTER 2

# ACCOUNTABILITY AND TEACHING PRACTICES: SCHOOL-LEVEL ACTIONS AND TEACHER RESPONSES

Laura S. Hamilton, Brian M. Stecher,  
Jennifer Lin Russell, Julie A. Marsh and Jeremy Miles

Policymakers and school reformers continue to look to standards-based accountability (SBA) policies as a vehicle to promote improvements in teaching and learning. This policy strategy gained traction nationwide – after a decade of state-led efforts – with the passage of the federal No Child Left Behind (NCLB) Act of 2001, which mandated that all states enact a system of curricular standards and testing, yearly performance targets, and incentives. The popularity of SBA as a reform strategy stems from a belief that alignment of goals and imposition of incentives are necessary to overcome educators’ presumed resistance to improving instruction (Cohen, 1996). However, policymakers’ optimism should be tempered by several decades of research that has emphasized the limited governability of teaching by centralized actors or rules (Dreeben, 1970; Lortie, 1975). The pervasive images of classrooms decoupled from administrators and the environment (Meyer & Rowan, 1977, 1978; Weick, 1976) and strong occupational norms of teacher autonomy (Goodlad, 1984; Little, 1990;

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
**Research in Sociology of Education, Volume 16, 31–66**  
Copyright © 2008 by Emerald Group Publishing Limited  
All rights of reproduction in any form reserved  
ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16002-5

Lortie, 1975) portray considerable barriers to widespread change inside classrooms.

The adoption of policies that involve systemic reform and externally imposed incentives has been one response to these challenges. A growing body of research on standards-based reform and high-stakes testing conducted over the past decade suggests these policies do influence teachers' instructional practices, which raises the question of how SBA may have helped to overcome the challenges associated with coordinating educational improvement efforts from higher levels of the education hierarchy. One of the distinctive features of current SBA policies that may explain their local effectiveness is a focus on the school as the unit of accountability. Test results are published for each school, and schools are evaluated against the state's Adequate Yearly Progress (AYP) targets. In addition, the interventions specified in the law occur at the school level rather than being focused on specific teachers. As we discuss below, changes to classroom instruction are hypothesized to occur in large part as a result of school-level responses to SBA.

We examine the extent to which teachers' responses to SBA policies reflect school-level conditions and actions. We draw on a larger longitudinal study, the Implementing Standards-Based Accountability (ISBA) project, which is examining how district staff, principals, and teachers in three states (California, Georgia, and Pennsylvania) are responding to the SBA provisions under NCLB during a three-year period that includes the 2003–2004, 2004–2005, and 2005–2006 school years (see Stecher, Hamilton, & Naftel, 2005; Hamilton et al., 2007). This chapter presents descriptive data from the third year of the study with a focus on classroom-level responses to SBA among mathematics teachers. We also present preliminary efforts to model the relationship between teacher responses and actions at the school level.

We find that accountability mandates are influencing teachers' instructional practices in distinct ways, but that variation in how teachers respond to SBA policy is primarily found among teachers within schools rather than between schools or districts. Schools and districts do not appear to have strong effects on how teachers orient their instructional practices toward assessments or use assessment data to guide instruction (two actions central to the SBA theory of action). In this light, we argue that the lack of influence of schools and districts on teachers' instructional adaptations suggests that SBA has not overcome the organizational challenges to coordinated instructional improvement. Yet while policy operates directly on teachers, the mechanisms through which that impact occurs are not well understood.

## BACKGROUND

The design of the ISBA project was guided by an analysis of the SBA theory of action, its likely effect on educators' work across levels of the educational hierarchy, and prior research on the impact of SBA policies on teachers' work. We begin placing our work in the context of theoretical accounts of school organizations and the occupational norms of teaching.

### *Perspectives on Organizational Reform*

Sociological accounts of schools as organizations and teaching as an occupation draw attention to the particular challenges associated with coordinating educational improvement efforts mounted from higher levels of the education hierarchy. The prevailing notion of the organizational structure of public schooling has portrayed classrooms as decoupled from their school context and broader environment. Organizational theorists have long argued that schools and classrooms are loosely coupled with higher levels. Schools respond to pressures from their environment by making largely symbolic changes to organizational structures and buffering classroom practice from change and outside scrutiny (Firestone, 1985; Meyer & Rowan, 1977, 1978; Weick, 1976). Any close monitoring of instruction potentially exposes problems that undermine public trust in schools, so the educational bureaucracy has historically attended to institutionalized rules defining the structure of schooling, and paid far less attention to the technical action of teaching found inside classrooms (Rowan & Miskel, 1999). As a result, teachers retained considerable control over instructional practice in their largely isolated classrooms.

Sociologists emphasize that occupational norms and workplace conditions further impede external control of instructional practice. The goals of teaching are considered ambiguous, multiple, and variable, reflecting the political decentralization of American education (Rowan, 1990, 1995). The nature of the teaching task is said to demand significant teacher autonomy: teaching "is intractable to organizational routines because it involves an artful balance of universalistic evaluation and particularistic motivation that requires continuing infusions from a teachers' own fund of classroom experience, and consequentially, substantial classroom autonomy" (Bidwell, 2001, p. 103). Teaching demands flexibility to meet the varied needs of children and gain the cooperation of students to engage in academic tasks (Dreeben, 1970). And

teaching as an occupation is subject to uncertainty, vulnerability, and strong norms of privacy and non-interference (Little, 1990; Lortie, 1975). Other explanations for why teaching is difficult to manage at the school level include the ideology of teacher professionalism, which tempers hierarchical oversight over teaching (Firestone, 2003); zones of authority in schools which define the classroom as the teacher's domain, while schoolwide issues are the purview of principals (Ingersoll, 2003; Marks & Louis, 1997); and the inherent complexity of classroom supervision (Dreeben, 1970).

While theoretical accounts of the organizational structure of U.S. schools have long emphasized these barriers to coordination to linking classroom practice to higher levels of authority in the educational system, some observers note that the institutional environment shaping public education is changing: there is "greater emphasis on monitoring organizational performance" and "a growing attempt to develop more coherent educational policy" (Rowan & Miskel, 1999, p. 379). Smith and O'Day's (1991) call for systemic reform argued that a "fundamental barrier to developing and sustaining successful schools in the USA is the fragmented, complex, multi-layered educational policy system in which they are embedded" (p. 237). The systemic reform strategy aimed to overcome loosely coupled organizational structures through state-led educational reform emphasizing unified goals, a coherent system of instructional guidance, and restructured educational governance. Some of the concrete manifestations of this approach have included school-level efforts to coordinate, support, and monitor instruction. These efforts include emphasizing principals' roles as instructional leaders, promoting mentoring among teachers, and instituting coaching models for teacher improvement. Organizational theories predict that the shift toward systemic reform would lead to tighter linkages among goals, activities, and outcomes, but also suggest that powerful forces must be overcome for this shift to occur (Rowan & Miskel, 1999). For example, the deeply-ingrained tradition of local control in education and occupational norms of autonomy and non-interference in teaching challenge the efficacy of systemic reform efforts.

### *Standards-Based Accountability in Theory*

In theory, SBA systems have the potential to overcome some of the barriers represented by the fragmented nature of the educational system and to have a substantial influence on teaching. NCLB works to improve education by

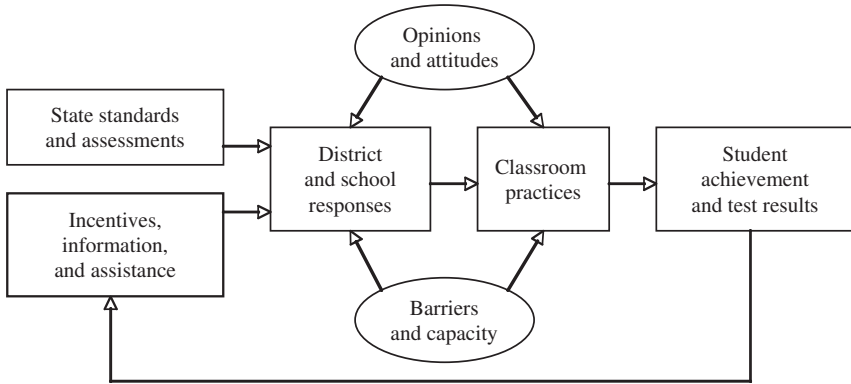


Fig. 1. Standards-Based Accountability Theory of Action.

defining desired student learning outcomes and attaching incentives to aligned measures of student performance. In simple terms, the SBA theory of action can be described as a feedback loop, as illustrated in Fig. 1. The states adopt a set of standards that describe what students should know and be able to do, as well as assessments to measure student mastery of the standards. Districts and schools implement curriculum and instruction practices that teach the skills embodied in the standards. Incentives for schools and districts are established for student performance on assessments. These incentives could include rewards for positive results and interventions for negative results. Incentives, along with assistance and support that might accompany them, cause educators to eliminate ineffective practices while reinforcing effective ones. (See Stecher, Hamilton, & Gonzalez, 2003 for a more detailed discussion of the assumptions of SBA under NCLB.)

Under the NCLB Act, schools are the primary unit of accountability. Test scores are aggregated to the school level and incentives apply initially to schools (rather than individual teachers or students). For a number of practical and historical reasons, this feedback cycle usually occurs annually, with testing at the end of the year leading to consequences in the following year. The Congress adopted NCLB because they believed this approach to accountability would increase educators’ attention to student learning – specifically the learning goals embodied in the state standards. But we know surprisingly little about how the school acts to mediate accountability pressures for teachers.

Implicit in its theory of action are a number of mechanisms through which SBA might lead to improved (or at least better aligned) instruction and improved student achievement. First, these systems may create *incentives* (e.g., rewards and sanctions tied to test results) that are designed to motivate teachers to focus on tested material. Second, they may provide teachers and other educators with more *information*, in the form of test results, to help them evaluate which students and which content areas need extra attention. Finally, most SBA systems include an *assistance* component that is intended to provide low-performing schools with extra help and guidance. All of these mechanisms have the potential to overcome the organizational challenges stemming from the loose coupling and lack of alignment discussed earlier.

Fig. 1 also illustrates the fact that, although schools are the primary unit of accountability in the eyes of policymakers, the responsibilities for carrying out SBA provisions are distributed across levels. Under NCLB there is a cascade of responsibility from the higher to the lower levels of the system. The state sets standards and adopts assessments, districts must intervene if schools fail to meet them, school leaders are responsible for motivating teachers and improving their practice, and teachers respond to these layers of incentives and exhortations when designing their instruction. Responses to NCLB are also likely to flow in a top-down manner. District actions should influence school actions, which in turn should affect what teachers do in the classroom. These effects may build over time to promote changes (positive or negative) in student achievement. In addition, these cross-level relationships are likely to be mediated by educators' attitudes and opinions about SBA and by various aspects of schooling context (e.g., capacity, knowledge, skills, and resources) that can act as barriers or enablers to successful implementation. With the implementation of NCLB distributed across levels of the educational hierarchy and incentives for district and school staff to alter practices to align with accountability systems, SBA systems have the potential to alter prevailing organizational and occupational conditions.

### *Standards-Based Accountability in Practice*

An emerging body of evidence is illuminating the effects of NCLB-mandated SBA provisions on districts, schools, and teachers. In some cases, NCLB seems to influence teachers directly; in other cases, the law appears to operate through school administrators to influence teachers indirectly. The

degree to which NCLB touches the work of teachers directly, or operates indirectly, through school-level responses and administrators' actions, provides some insight into the mechanisms through which accountability policies are altering organizational and occupational barriers to coordinated classroom improvements.

Some evidence suggests that SBA policies have led to direct changes in teacher practice that are generally considered beneficial for improving student learning. For example, according to district and school administrators, NCLB appears to be raising learning expectations, focusing attention on traditionally low-performing groups, promoting greater alignment between standards and instruction, and increasing the use of data for decision-making (Center on Education Policy, 2006; Hamilton et al., 2007). These responses are consistent with the theory of action underlying SBA, and with the hopes of NCLB's authors and advocates.

On the other hand, some teacher responses to SBA may not be beneficial for student learning. For example, many teachers report narrowing their instruction within the tested subjects to focus on tested content and de-emphasize content that is not tested (Hamilton et al., 2007). These findings are consistent with those from earlier studies of high-stakes testing which identified changes including shifts in emphasis among tested and non-tested subjects (Jones et al., 1999; Koretz, Barron, Mitchell, & Stecher, 1996; Shepard & Dougherty, 1991; Smith, Edelsky, Draper, Rottenberg, & Cherland, 1991; Stecher, Barron, Chun, & Ross, 2000; Hamilton, 2003), changes in the timing of topic introduction to match the testing schedule (Corbett & Wilson, 1991; Darling-Hammond & Wise, 1985), shifts in emphasis among tested and non-tested content within a subject (Shepard & Dougherty, 1991; Romberg, Zarinia, & Williams, 1989; Smith et al., 1991), and use of instructional materials designed to mirror the state test (Pedulla et al., 2003). This literature indicates that high-stakes testing exerts a powerful effect on teachers' practices, but to date most of this work has been conducted on relatively small or unrepresentative samples of teachers, and none of it examines teachers' responses to the specific features of the NCLB legislation.

In addition to effects of accountability systems that are directly experienced by teachers, research suggests that NCLB has effects at the school level that influence teachers jointly and/or through changes in leadership. For example, some schools report responding to NCLB by reducing or eliminating instruction in non-tested subjects such as art, physical education, and foreign languages (Dobbs, 2004). The degree to which schools can mobilize collective responses to external accountability may be a factor in whether schools improve student achievement in response



to SBA policies (Abelmann, Elmore, Even, Kenyon, & Marshall, 1999). Research suggests that principals can play an important role in shaping teacher practice (Kirby, Berends, & Naftel, 2001; Knapp, Copland, & Talbert, 2003) and promoting positive changes in instruction. For example, a large body of literature has shown that principals' instructional leadership is related to the likelihood of school change and student learning (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Waters, Marzano, & McNulty, 2003) and that professional development efforts can result in changes in teachers' knowledge and beliefs as well as in their instructional practice (Loucks-Horsley & Matsumoto, 1999).<sup>1</sup> In particular, principals' support for instruction aligned with standards and assessments has been found to contribute to changes in the amount of test-focused instruction and inquiry-oriented instruction in which teachers engage (Firestone et al., 2004). In low-capacity schools, principals who are able to mobilize organizational resources to change school operations are more successful in raising student achievement in response to external accountability (Lemons, Luschei, & Siskin, 2003).

In sum, prior research suggests that SBA policies may be influencing teachers' practice, despite longstanding organizational and occupational conditions that have historically buffered instruction from external control. Evidence of direct influences on teachers' instructional practices suggests that policies can penetrate classroom boundaries. In addition, the research shows that schools as organizations can likely play a role in providing support and encouraging positive instructional responses to accountability mandates but that this role does not come naturally and represents a dramatic change in the typically decoupled relationship between school administration and classroom instruction. This chapter examines the ways in which teachers are changing their practices as a result of NCLB and the extent to which these changes are consistent within schools. We focus on behaviors relating to standards, assessments, and assessment use because they are core elements of the law.

## RESEARCH QUESTIONS

Our study focuses on three fundamental empirical questions:

- What kinds of practices related to SBA, such as the use of data and test-preparation activities, are reported by elementary and middle schools mathematics teachers?

- To what extent do mathematics teachers working in the same schools report similar practices?
- What school-level actions and conditions are associated with teachers' mathematics instructional practices?

These questions are descriptive rather than causal, because NCLB was implemented universally without a true experimental design. The first question is intended to provide information about teachers' pedagogy and the extent to which their practices are consistent with what would be expected in the context of SBA several years after NCLB was enacted.<sup>2</sup> We report separately for elementary and middle school teachers, because while both groups face significant testing requirements under NCLB, the different contexts in which elementary and middle school teachers work might influence their responses to the policy. The second and third questions are motivated by the idea that teachers' practices are likely to be influenced by school-level actions. Questions two and three also provide insight into the degree to which SBA policies enable tighter connections between school administrators' actions and teachers' instructional practices. All three questions address the extent to which SBA policies are penetrating classroom boundaries.

## **DATA AND METHODS**

Detailed descriptions of the study methods are provided elsewhere (Stecher et al., 2006; Hamilton et al., 2007). These sources also describe the SBA systems and the broader policy environments in California, Georgia, and Pennsylvania. In summary, we selected these three states to represent the range of approaches to implementing SBA under NCLB and to provide both geographic and demographic diversity. The states differ in a number of ways that might affect their implementation of SBA, including size of their K-12 systems, the diversity of student population, and the educational policy context. California – the largest and most diverse of the three states – had considerable experience with its own SBA system before NCLB. Much smaller and less diverse than California, Georgia had begun to implement SBA for the first time just before NCLB was enacted. Pennsylvania, which enrolls the least diverse student population of the three, is a state with a tradition of local control and had just begun to use state tests for accountability on a limited basis prior to the inception of NCLB. In the next section, we briefly describe our surveys and analytic approach.

*Surveys*

The results presented in this chapter rely primarily on teacher survey data gathered from state-representative samples of elementary and middle school mathematics teachers.<sup>3</sup> The numbers of respondents for the 2005–2006 survey administration are provided in [Table 1](#); these numbers are similar to those obtained in the previous two waves of data collection. The overall response rates in 2005–2006 were 73 percent for superintendents, 86 percent for principals, and 82 percent for teachers. We generated state-specific sampling and non-response weights and used these to produce the descriptive estimates. Thus, the statistics in the descriptive tables that follow are the estimates of the responses of teachers representative of each state – California, Georgia, and Pennsylvania. The project also administered surveys to representative samples of principals and superintendents. The principal survey data are used in the modeling results presented in the final section. Details on data collection for these groups can be found in [Hamilton et al. \(2007\)](#).

The few schools that included both elementary and middle-level grades (e.g., K-8 schools) are included in both the middle and the elementary school samples, but the estimates reported here are based only on teachers in the relevant grades (e.g., the middle school estimates include only teachers in grades 7 and 8). Some of the teacher survey questions focused specifically on mathematics; for these questions, only teachers who taught the relevant subject responded.

Because our surveys were designed to address a broad range of topics related to testing and accountability, we could allocate only limited numbers of items to each of the research questions described in the earlier section. It is important to acknowledge that the findings reported in this chapter rely on self-reported data. Some prior research suggests that well-designed surveys can measure teachers' instructional practices with a reasonable degree of accuracy ([Le et al., 2006](#); [Mayer, 1999](#); [Mullens & Gayler, 1999](#);

**Table 1.** Numbers of Survey Respondents for Each State, 2005–2006.

	District Superintendents	Principals	Elementary School Math Teachers	Middle School Math Teachers
California	26	75	375	187
Georgia	24	96	627	428
Pennsylvania	20	88	565	266
Total	70	259	1567	877

Smithson & Porter, 1994). A study by Hill (2005) did cast some doubt on the validity of teachers' responses to survey questions addressing the mathematical content of their lessons, which seems to reflect the inconsistency among teachers in how they define terms and describe content. The surveys used in the ISBA study do not include content-specific items, however, so Hill's findings are not of great concern here.

For many of the constructs the surveys were designed to measure, we were able to draw from existing survey items that have been administered and validated in a variety of contexts. Scales measuring constructs such as principal leadership, for example, have been used and found to function well in a number of studies. The items addressing specific instructional practices in response to state tests were developed for a pilot study of teachers in Massachusetts and were subjected to extensive cognitive interviews and other validation (Koretz & Hamilton, 2003). In some cases, particularly for items asking directly about NCLB or specific state policies, it was necessary to develop new items. All items were tested through cognitive interviews with a small number of teachers. In addition, psychometric analyses were conducted after the 2004 survey administration and were used to inform decisions about which items to retain in subsequent waves.

Moreover, it is important to note that the study was not designed to explain differences across the three states; instead, the states should be considered three case studies of states that were at different points in their experiences with accountability before NCLB was implemented. We report results separately by state because our sampling approach was designed to provide samples that were representative of each state rather than to support inferences about responses across the combined set of states.<sup>4</sup>

### *Analytic Approach*

As discussed earlier, the first set of descriptive results are based on weighted survey responses that provide state-representative estimates. We address the second research question by examining the proportion of variability in teacher responses that occurs within versus between schools and districts. We address the third research question by estimating a set of regression models within a multilevel modeling framework to examine relationships between measures of teacher and principal responses to SBA. Additional details regarding the modeling methodology are provided in the section where the modeling results are presented.

## RESULTS

In the following sections, we focus on a selected set of findings related to each of the three research questions. First, we describe overall patterns of practice related to standards and assessments. Then, we explore the influence of schools on practice by analyzing the consistency of practice within schools. Finally, we examine associations between specific administrative practices and teacher behaviors. In some cases, we present data from both elementary and middle school teachers, particularly when their responses differed. In other cases, we focus on elementary teachers, because they generally have more discretion with respect to changes in practice. For example, they can more easily change the amount of time allocated to content area instruction.

### *Instructional Practices Related to SBA*

The descriptive results presented in this section provide information about California, Georgia, and Pennsylvania mathematics teachers' practices during the 2005–2006 school year. We focus on instructional practices that are likely to be affected by SBA, and we also present a few selected findings related to teachers' attitudes toward their state SBA systems. Together, these results illustrate a variety of ways that teachers' practices appear to be shaped by their states' SBA policies. Although the study design does not support causal attribution, teachers' responses strongly suggest that state standards and assessments have directly influenced their day-to-day work.

### *Instructional Time*

Many critics of SBA have expressed concern about a reduction of time devoted to subjects that are not tested, and a number of districts have reported making such scheduling changes ([Center on Education Policy, 2006](#)). Given that elementary teachers in self-contained classrooms have more flexibility to change the amount of time spent on different subjects than do teachers who are assigned to teach specific classes during specific periods (a group that includes most middle school teachers), we focus on elementary teachers here. As shown in [Table 2](#), although the reports of elementary teachers in our study provide some indication of changes in time allocation, in every case a majority of teachers reported no change. The percentages of teachers reporting an increase in instructional time were largest for mathematics and reading. Time spent on science and social

**Table 2.** Percentage of Elementary School Teachers Reporting Changes in Instructional Time from 2004–2005 to 2005–2006.

	California			Georgia			Pennsylvania		
	Decrease	No change	Increase	Decrease	No change	Increase	Decrease	No change	Increase
Mathematics	4 (1)	71 (4)	19 (3)	6 (1)	68 (3)	35 (2)	4 (1)	65 (4)	26 (3)
Science	19 (3)	60 (6)	15 (3)	12 (2)	67 (2)	7 (2)	18 (3)	61 (3)	11 (3)
Reading/ Language Arts/ English	3 (1)	59 (6)	31 (6)	4 (1)	63 (3)	33 (3)	3 (1)	66 (3)	23 (3)
Social Studies	23 (4)	62 (5)	9 (2)	11 (2)	67 (2)	6 (2)	22 (3)	61 (3)	6 (1)
Arts/Music	23 (4)	61 (5)	9 (2)	6 (2)	70 (3)	1 (2)	1 (1)	89 (1)	1 (0)
Physical education	15 (3)	70 (5)	9 (2)	6 (2)	77 (2)	1 (1)	2 (2)	86 (2)	3 (1)

*Notes:* Teachers were asked whether the instructional time provided to their students in each subject stayed the same, increased, or decreased between 2004–2005 and 2005–2006. Omitted category is “don’t know.” Standard errors are shown in parentheses.

studies was more likely to decrease than increase, and in several cases time in these subjects was more likely than physical education or arts to decrease. These changes build on changes of similar magnitude measured in the previous two years of the study, which suggests that the total change occurring as a result of NCLB is likely to be larger than what is observed for a single year. It is worth noting that both California and Georgia had state science assessments in place, and Pennsylvania was piloting one in preparation for the NCLB science-testing requirement, but none of the science results affected schools’ AYP status. The exclusion of science from AYP probably accounts for the relatively infrequent reports of increased time spent on science instruction. To the extent that these changes can be attributed to NCLB, these results suggest that SBA has exerted an influence on teaching in the most fundamental sense, by increasing the amount of time teachers spend on specific subjects. However, instructional time is an aspect of practice that is relatively easy to change as a result of policies and practices adopted by higher levels of the education system (e.g., through district-wide scheduling mandates). The content and style of instruction are much more resistant to such changes, as discussed earlier; the next several sets of findings address these outcomes.

#### *Alignment with Curriculum and Student Assessments*

As noted in the introduction, high-stakes testing has been shown to lead not only to changes in allocation of time across tested and non-tested subjects,

**Table 3.** Percentage of Teachers in Tested Grades Who Report Aligning Their Instruction with State Content Standards and State Assessments, 2006.

	California		Georgia		Pennsylvania	
	Elementary	Middle	Elementary	Middle	Elementary	Middle
Mathematics content standards	94 (1)	95 (2)	86 (2)	88 (2)	84 (3)	80 (4)
Mathematics assessments	58 (4)	53 (6)	78 (3)	80 (3)	85 (3)	85 (4)

*Notes:* Standard errors are in parentheses. Response options included strongly disagree, disagree, agree, strongly agree, and I don't know; table entries are percentage agreeing or strongly agreeing. In Pennsylvania, this question focused on the Assessment Anchors rather than the Pennsylvania Academic Standards, because our state contacts told us that teachers were encouraged to use the Anchors as a source of information about what was expected of them.

but also to reallocation of emphasis on tested and non-tested content within a subject area. Content standards provide one mechanism for promoting these types of changes. They are intended to help teachers determine what topics and skills should be prioritized, and which are most likely to be included on state tests. The tests themselves also send signals about what content to emphasize, and studies of alignment have shown that the tests and standards do not always send consistent messages (Rothman, Slattery, Vranek, & Resnick, 2002). Table 3 shows that large majorities of mathematics teachers reported aligning their instruction with the state standards. In Georgia and Pennsylvania, the proportions who reported aligning instruction with state assessments were very similar to the proportion aligning with standards, but the former numbers were significantly lower than the latter in California. Based on conversations with state and district officials in the three states, it appears that one reason for this difference is that California educators have less access to released test items and other test-preparation materials than do educators in the other two states (see Hamilton et al., 2007 for additional discussion of the state contexts).

As discussed elsewhere (Hamilton et al., 2007), districts and schools took a number of steps to promote alignment of instruction with standards, such as providing pacing guides that helped ensure coverage of the standards throughout the year. At the same time, teachers expressed concern about insufficient alignment between state standards and tests and between state standards and their local curriculum, and described a number of challenges associated with efforts to align instruction to standards. For example, large

majorities of teachers in each state reported that their state content standards included more material than could adequately be covered in a year, and several teachers we interviewed said they felt the need to rush through the curriculum and often moved onto a new topic before students had mastered the previous one because of the need to cover so much material. This is clear evidence that state-level mandates have worked their way into the classroom.

### *Instructional Practices Attributed to Testing*

The next set of results examines teachers' perceptions of how their instruction has been affected by state accountability tests. We asked a set of questions similar to one developed by [Koretz and Hamilton \(2003\)](#) to understand teachers' responses to high-stakes testing. Teachers were asked to describe the degree to which their teaching was affected by the state test. Specifically, we asked teachers to, "Think about ways in which your teaching is different because of the [state test]<sup>5</sup> than it would be without the [state test]. How much do the following statements describe differences in your teaching due to the [state test]?" Teachers could select "no difference," "differs by a small amount," "differs by a moderate amount," or "differs by a great deal."

The results for elementary teachers are presented in [Table 4](#). Teachers report a variety of ways in which their practices are influenced by state assessments. Some of the more prevalent responses represent SBA effects that are often viewed as desirable. These include searching for more effective teaching methods and spending more time teaching content. Others are practices that have been associated with excessive curriculum narrowing and a risk for score inflation – in particular, emphasizing specific problem styles and formats and teaching test-taking strategies. In addition, between 30 and 41 percent of elementary teachers report focusing more on students who are close to proficient (sometimes called "bubble kids"; see, e.g., [Booher-Jennings, 2005](#)). Although we do not know whether this is done at the expense of other students, it does suggest a possible reason to be concerned about the instruction provided to students scoring far below or above the level of proficiency, a concern that is echoed by teachers as described in a later section of this chapter.

The responses of middle school teachers ([Table 5](#)) were similar to those for elementary teachers, with a few exceptions. In all three states, middle school teachers were much more likely than elementary teachers to report offering assistance to students outside of school. Middle school teachers also



**Table 4.** Percentage of Elementary School Teachers Reporting Their Instruction is Different as a Result of Mathematics Assessments, 2006.

	California	Georgia	Pennsylvania
Assign more homework	31 (4)	32 (3)	30 (3)
Search for more effective teaching methods	62 (5)	76 (3)	60 (3)
Focus more on standards	68 (5)	76 (3)	78 (3)
Focus more on topics emphasized in assessment	64 (4)	71 (4)	71 (4)
Emphasize assessment styles and formats of problems	56 (5)	75 (3)	81 (3)
Spend more time teaching test-taking strategies	53 (5)	59 (3)	56 (3)
Spend more time teaching content	51 (5)	58 (3)	55 (4)
Focus more on students who are close to proficient	41 (4)	39 (3)	30 (3)
Offer more assistance outside of school for students who are not proficient	27 (4)	33 (4)	21 (3)
Rely more heavily on multiple-choice tests	29 (5)	37 (3)	23 (2)
Rely more heavily on open-ended tests	21 (3)	27 (2)	46 (3)

*Notes:* Standard errors are provided in parentheses. Response options were: not at all, a small amount, a moderate amount, and a great deal. Cell entries are percentages reporting they engage in each practice “a moderate amount” or a “great deal” as a result of the state tests. These questions were not presented to PA science teachers because of the lack of a statewide science test.

**Table 5.** Percentage of Middle School Mathematics Teachers Reporting Their Instruction is Different as a Result of Mathematics Assessments, 2006.

	California	Georgia	Pennsylvania
Assign more homework	21 (4)	23 (2)	28 (4)
Search for more effective teaching methods	56 (5)	72 (2)	75 (5)
Focus more on standards	52 (6)	69 (3)	72 (5)
Focus more on topics emphasized in assessment	45 (6)	48 (3)	41 (4)
Emphasize assessment styles and formats of problems	40 (4)	53 (2)	49 (5)
Spend more time teaching test-taking strategies	22 (3)	39 (4)	31 (3)
Spend more time teaching content	57 (5)	69 (3)	59 (6)
Focus more on students who are close to proficient	30 (4)	44 (4)	22 (4)
Offer more assistance outside of school for students who are not proficient	69 (5)	72 (3)	74 (4)
Rely more heavily on multiple-choice tests	19 (6)	35 (5)	17 (4)
Rely more heavily on open-ended tests	17 (5)	20 (3)	46 (6)

*Notes:* Standard errors are provided in parentheses. Response options were: not at all, a small amount, a moderate amount, and a great deal. Cell entries are percentages reporting they engage in each practice “a moderate amount” or a “great deal” as a result of the state tests.

tended to be less likely than elementary teachers to report emphasizing test-taking strategies, specific problem formats, and tested topics.

*Use of Achievement Data for Instructional Decision-Making*

SBA policies emphasize the importance of using information on student achievement to make instructional decisions. As discussed earlier, districts and schools have taken a number of steps to promote data-driven decision-making among their teachers. The next set of results summarizes responses to questions about the frequency with which teachers used various data-related practices throughout the year. In contrast to the results presented in Table 5, these questions did not ask teachers to attribute the practices to the state test, so they simply provide a descriptive picture of strategies used by mathematics teachers in the three states. As shown in Table 6, which indicates the percentage of teachers who reported using a technique “sometimes” or “often,” certain strategies were nearly universal among elementary teachers. Almost all teachers reported re-teaching content in response to assessment or assignment performance and reviewing assessment results to make decisions about course content or individual students. Middle school teachers were slightly less likely than elementary teachers to

**Table 6.** Percentage of Elementary School Mathematics Teachers Reporting Use of Data for Instructional Decision Making, 2006.

	California		Georgia		Pennsylvania	
	Elementary	Middle	Elementary	Middle	Elementary	Middle
Plan different assignments based on performance	80 (3)	67 (5)	87 (2)	83 (2)	79 (3)	73 (5)
Re-teach topics because performance on assignments or assessments did not meet expectations	97 (1)	97 (1)	98 (1)	95 (1)	95 (1)	90 (3)
Review assessment results to identify individual students who need supplemental instruction	94 (2)	91 (3)	96 (1)	91 (1)	93 (1)	80 (5)
Review assessment results to identify topics requiring more or less emphasis in instruction	93 (2)	86 (2)	96 (1)	95 (1)	91 (1)	81 (5)
Conduct a pre-assessment to find out what students know about a topic	67 (3)	51 (5)	76 (3)	68 (3)	58 (3)	56 (10)

*Notes:* Standard errors are provided in parentheses. Response options were: never, rarely, sometimes, and often. Cell entries are percentages reporting they engage in each practice “sometimes” or “often.”

report using these practices, but most of them were widespread at the middle school level, especially in Georgia. This pattern of responses could reflect similarities in teachers' responses due to mediating actions by districts or schools, or they could be common but independent reactions by individual teachers. While it is clear that SBA is penetrating classroom boundaries, it is unclear whether these effects also represent evidence of a tighter connection between school administration and teacher practice.

### *Opinions About State Tests and Accountability Systems*

The widespread use of achievement data described earlier is one of the responses that SBA advocates typically hope to promote. Teachers have access to a variety of sources of information about achievement, including the state test and the tests and assignments they use on a daily basis. We were particularly interested in understanding the extent to which teachers viewed the state accountability test as useful for making instructional decisions. Table 7 shows the percentage of elementary and middle school teachers agreeing with statements about the utility of these tests. It suggests fairly large differences among states, with Georgia teachers responding more favorably than teachers in the other two states. Moreover, similar percentages of Georgia teachers responded positively to each of the three items. In California and Pennsylvania, by contrast, teachers were less likely to report that the tests helped them tailor instruction to individual student needs compared with the other two uses about which we asked.

**Table 7.** Percentage of Mathematics Teachers Agreeing with Statements about the State Tests, 2006.

State Test Results	California		Georgia		Pennsylvania	
	Elementary	Middle	Elementary	Middle	Elementary	Middle
Allowed me to identify areas where I need to strengthen my content knowledge or teaching skills	73 (3)	62 (5)	89 (2)	79 (2)	62 (4)	57 (5)
Helped me identify and correct gaps in curriculum and instruction	67 (3)	60 (6)	88 (2)	80 (3)	56 (4)	61 (4)
Helped me tailor instruction to individual student needs	59 (3)	42 (7)	85 (2)	79 (3)	36 (3)	42 (6)

*Notes:* Teachers who said they did not receive test results are excluded. Response options were: strongly disagree, disagree, agree, and strongly agree. Standard errors are provided in parentheses.

These results are consistent with findings from the larger study showing more positive attitudes toward and greater use of data among educators at all levels in Georgia. For example, educators in Georgia were more likely to report the use of progress tests and more likely to express positive views about the usefulness of state and local test data than their counterparts in the other two states. These differences may be due to a related trend observed in our data: Georgia appears to be providing more overall support to schools and districts. Georgia administrators at the district and school levels were more likely than their counterparts in the other states to report providing and receiving needed assistance in a wide range of areas, including data use.

The final set of questions we include here examines teachers' opinions about the perceived effects of SBA. These results (Table 8) suggest that the hopes of some accountability advocates are being realized. In particular,

**Table 8.** Percentages of Elementary and Middle School Teachers Agreeing with Statements about State Accountability System, 2006.

	California		Georgia		Pennsylvania	
	Elementary	Middle	Elementary	Middle	Elementary	Middle
The state's accountability system supports my personal approach to teaching and learning	33 (5)	33 (4)	53 (3)	47 (3)	27 (3)	28 (5)
The state's accountability system leaves little time to teach content not on the state tests	89 (2)	88 (3)	82 (2)	83 (2)	87 (2)	95 (2)
Because of pressure to meet the AYP target, I am focusing more on improving student achievement at my school	78 (3)	72 (4)	84 (2)	81 (2)	80 (4)	78 (4)
The possibility of my school receiving rewards or sanctions is a very strong motivator for me	35 (5)	31 (4)	57 (3)	50 (3)	36 (3)	38 (4)
Overall, the state's accountability system has been beneficial for students at my school	37 (5)	34 (5)	54 (3)	50 (2)	31 (4)	28 (3)
As a result of the state accountability system, high-achieving students are not receiving appropriately challenging curriculum or instruction	52 (4)	54 (3)	46 (3)	51 (2)	46 (3)	56 (5)

*Notes:* Standard errors are provided in parentheses. Response options were: strongly disagree, disagree, agree, and strongly agree. Percentages in table are those responding agree or strongly agree.

large majorities of teachers in all three states report focusing more on student achievement because of pressure to meet the AYP target. But there is also reason for concern. Even larger majorities said that they had little time to teach content not on state tests – a finding that might be viewed positively by those who believe that state tests include all of the material that should be taught in schools, but one that is worrisome for those who believe such tests are limited in their scope. Approximately half of Georgia teachers reported that the accountability system had been beneficial for their students overall; the numbers were smaller in the other two states. The same pattern was observed for views about whether the accountability system supports teachers' personal approach to teaching and learning. Finally, approximately half of the teachers in all three states and at both levels endorsed the statement that because of the accountability system, high-achieving students are not receiving appropriately challenging curriculum and instruction. This last finding is consistent with district and school reports of instructional support activities focused on low-performing students (see [Hamilton et al., 2007](#)) and with teachers' reports of attention to "bubble kids." Clearly teachers are making changes that are inconsistent with their own values, which suggests that these policies have in fact reduced teachers' autonomy and control over their classrooms.

In summary, the results reported in this section illustrate that SBA policies affect the daily practice of teachers in classrooms. However, the nature of the effects of state policies on teachers' practices may also be influenced by what happens at the district and school levels. In the next section, we examine the extent to which teachers' responses to selected survey questions are similar within schools and districts, which provides insight into whether SBA policies directly affect teachers' practice or are the result of tighter coordination within schools and districts.

#### *Consistency of Teacher Reports Within the Same Schools and Districts*

The school and the district are the key units of analysis for the purpose of accountability under NCLB. The framers of the law assigned responsibility to schools and districts to develop strategies for improving student achievement. For example, the initial step once a school is identified for improvement is the development of a school improvement plan, and the actions that schools are encouraged to take include aligning curriculum and instruction with state standards and providing professional development. Many of the actions schools and districts take are designed to make the

instructional program more consistent and coherent, so it is worthwhile to examine the extent to which teachers within the same school or district act in a consistent manner.

To examine consistency of teachers' responses, we used intra-class correlation analysis, which partitions the total variance in teacher responses into separate components representing variance in responses among teachers in the same school, variance in responses (on average) among schools in the same district, and variance in responses (on average) among districts.<sup>6</sup> It assesses how similar a teacher is likely to be to another randomly selected teacher in the same school, or how similar a school is likely to be to another randomly selected school in the same district. We would expect the relative variance at each level to depend on a number of factors, including the likely strength of district or school influence. For example, if each district implements a distinct practice (e.g., a specific set of curriculum materials) and does so consistently across all schools and teachers, we would expect that a relatively high proportion of the variance in teacher reports about these practices would occur among districts and a relatively low proportion would occur among schools within districts or among teachers within schools. In contrast, if teachers in a given school engaged in very different practices, and a similar range of practices was found in other schools, then the proportion of variance among teachers would be relatively high compared to the proportions for schools and districts.

Table 9 shows the proportion of variance for selected teacher-reported variables from the 2006 questionnaires. This is not an exhaustive list, but is intended to illustrate a range of teacher-level measures, including some practices that we would expect to have larger school-level components and other practices we would expect to vary greatly among teachers. We include a number of variables that are not related to standards and assessments for this reason. In a sense, they serve as validity check to see whether our analysis detects school-level variation in situations where we would expect to find it. Variables were selected to include some for which we would expect to find consistent practice among teachers in the same school and others for which this is not necessarily the case. As can be seen from the Table 9, perhaps the most striking finding is that on most items, teacher reports vary considerably within schools, but remain stable when comparing across schools and districts: The proportion of variance among teachers within schools is 0.70 or higher on most variables, while the proportions for schools and districts are usually below 0.20 and 0.10, respectively, on most variables.

Table 9 shows that there is little district- or school-level variability in teachers' opinions about the quality of their textbooks or about the validity

**Table 9.** Variance Decomposition for Selected Teacher Survey Responses, by State, 2005–2006.

	California		Georgia		Pennsylvania	
	Teacher	District	Teacher	School	Teacher	School
I am satisfied with the quality of mathematics textbooks and curriculum materials at my school	.90	.06	.91	.9	.83	.11
The (state) mathematics test is a good measure of students' mastery of the (state) standards in mathematics	.96	.01	.97	.03	.96	.03
Did your district or state establish a specific "pacing plan" ... to schedule instructional content throughout the year?	.64	.19	.79	.08	.64	.09
Are you required by your district or school to administer mathematics "progress" tests?	.58	.14	.45	.36	.45	.09
Teacher morale is high	.72	.26	.78	.21	.79	.17
The Principal communicates a clear academic vision for my school	.71	.29	.83	.17	.70	.30
I trust the Principal at his or her word	.74	.26	0	.77	.75	.25
I require students to have their parents sign-off on homework	.74	.26	0	.81	.68	.32
How often do you have students work on extended mathematics investigations or projects?	.87	.07	.94	.02	.95	.05
The possibility of my school receiving rewards or sanctions is a very strong motivator for me	.97	.02	.94	.05	.91	.05

of the state mathematics tests. Teachers seem to keep their own counsel with respect to these features of their educational systems. On the other hand, as we might expect, teachers within a district (or a school) are more likely to agree with each other when it comes to reports about district policies and actions, such as required “pacing plans” or progress testing. Between 14 and 47 percent of the variance in teachers’ responses to these items occurs at the district level (and another 10 to 30 percent occurs at the school level). These responses are consistent with the expectation that teachers within a school or district should have a shared understanding about such policies. Larger school variance components are also associated with questions about teacher morale, the characteristics of the principal, and homework policies, as might be expected.

Despite reports of district- and school-level efforts to promote consistent instruction, there is extensive teacher-level variation and very little school- or district-level variation in responses to questions about math teaching, such as whether the teacher assigns extended investigations. Similarly, the possibility of the school being singled out for a sanction or reward does not affect teachers in a consistent manner. There is almost no school-level variance on this item. Although not shown here, these results are consistent across most of the instructional practice and attitude measures discussed earlier.

It is also worth noting that proportions of variance among schools and districts vary among the states on some items. For example, required “progress tests” seem to be a district phenomenon in Pennsylvania but more of a school phenomenon in California and Georgia. Together, the results from the variance components analysis suggest an absence of strong effects of district or school policy and practice on teachers’ instruction and attitudes. Nonetheless, it is still worth exploring whether the relatively small proportion of between-school variance can be explained by school-level factors. The next section describes analyses designed to explore this relationship.

#### *Relationships Among Responses to SBA at the School and Classroom Levels*

Next, we describe the application of relatively simple multilevel analyses to examine the relationship between school- and classroom-level responses to SBA. Eventually, we will estimate more complex models to illuminate the relationships described in Fig. 1 using a longitudinal panel of data from



principals and teachers in the same schools. The models we describe here, which focus on exploring relationships between school- and classroom-level factors, serve as a starting point for these more complex analyses. Before presenting the modeling results, we briefly describe our measures and analytic approach.

### *Principal Measures*

We used our survey data to construct several principal- and teacher-level variables. The selection of principal variables for this exercise was informed by literature examining how instructional leadership influences teacher responses and student outcomes, and by earlier project findings regarding the strategies that school and district personnel viewed as most important for school improvement. Decisions about which of these measures to include in each model were based on a set of preliminary models that included each principal measure as a predictor of each of the two dependent variables, along with indicator variables for states and for school level (elementary versus middle school). Variables that showed significant relationships with the dependent variable were included in the full model for that dependent variable.

The first measure, *conditions promoting data use* or *data conditions*, measures principals' reports about a number of factors that are likely to promote effective use of achievement data for decision-making. It is a mean of the standardized versions of seven four-point Likert scale items (ranging from "strongly disagree" to "strongly agree") addressing contextual support for using data. An exploratory factor analysis of these items suggested a single scale, with an internal consistency reliability coefficient (alpha) of 0.77. The component items included:

- The test results we receive have explicit links to content standards and/or lesson plans.
- The district and/or state helps me understand the state accountability system requirements.
- The district and/or state helps my school staff understand the state accountability system requirements.
- The information we receive about our school's performance is clear and easy to understand.
- The information we receive about our school's performance is timely.
- Teachers in my school have the skills and knowledge needed to analyze and make use of the test results we receive.
- State test scores accurately reflect the achievement of students in my school.

The second measure, *AYP Pressure*, addresses principals' beliefs about how SBA pressure has influenced their schools. It consists of a single item ("Because of pressure to meet the AYP targets, my staff and I are focusing more on improving student achievement than we would without the AYP target") measured in the same four-point Likert scale ("strongly disagree" to "strongly agree").

*Instructional leadership* was assessed by asking principals about how much time they spent on a variety of activities directly related to instructional leadership (in contrast to management activities such as preparing budgets or responding to student discipline problems). Each item used a 1–4 scale (from 1 = no time spent on the activity; to 4 = over 15 hours in a given week). This *instructional leadership activities* scale is the standardized mean of the six items below ( $\alpha = 0.72$ ).

- Observing your teachers' classroom instruction.
- Providing feedback to teachers regarding curriculum and instruction.
- Reviewing student assessment results.
- Planning or conducting teacher professional development workshops.
- Meeting with school leadership teams to plan for school improvement.
- Collaborating with other principals.

Finally, the use of various school improvement strategies was assessed by asking principals whether they had engaged in each of a number of activities. Principals were asked to indicate which strategies they had used, and then to select the three strategies that they found most important for making their school better. After exploring several variables constructed from both sets of questions, we identified a small number of variables that showed consistent relationships with at least one of the dependent variables. We constructed indicators for whether the principal included each of the following as one of the three most important strategies: *increasing the use of student achievement data to inform instruction*, *increasing the quantity of teacher professional development*, and *using existing research to inform decisions about improvement strategies*. Other strategies that were commonly used and that seemed relevant to understanding influences on teachers' practices, such as efforts to match curriculum with state standards and assessments, were excluded because of lack of bivariate relationships with the outcomes of interest (probably due in part to low variance).

#### *Teacher Measures*

The first instructional practice measure, *test-focused instructional practices*, is constructed from the items presented in Tables 4 and 5. Although

these items were designed to assess different types of responses to state testing, an exploratory factor analysis suggested a single underlying construct measuring the extent to which teachers believe their practices have been influenced by the tests. The resulting scale has an internal consistency (alpha) of 0.89.

The second practice measure used items reported in Table 6. Exploratory factor analyses of these items suggested a three-factor solution, but only one of the three factors had sufficient variability and substantive coherence to be included in the modeling reported in this chapter. This factor, *use of achievement data*, included five items (alpha = 0.79):

- Plan different assignments or lessons for groups of students based on their performance on assignments or assessments.
- Re-teach topics because student performance on assignments or assessments did not meet your expectations.
- Review assessment results to identify individual students who need supplemental instruction.
- Review assessment results to identify topics requiring more or less emphasis in instruction.
- Conduct a pre-assessment to find out what students know about a topic.

#### *Multilevel Modeling Approach*

We estimated two multilevel models to explore relationships between school-level measures and the two measures of instructional practice described earlier. We estimated models with three levels: at level one were measurement occasions, with two occasions per school (change over the two years), level two was school, and level three was district. Thus we had measurement occasions nested within schools nested within districts. The predictors chosen for inclusion in each model differed as a result of the exploratory analyses discussed earlier and were all level one predictors. Both models included indicator variables for states, level (elementary versus middle), and for whether the school met its AYP target the previous year (“made AYP”). Year was included as a random effect, at the school level, and the intercept was random at both the school and district level. We tested interactions of state by year and school type by year. These were not statistically significant and therefore have not been reported.

#### *Modeling Results*

As a first step, we conducted a variance components analysis for each of the teacher dependent variables. As shown in Table 10 for 2005–2006, and

**Table 10.** Variance Decomposition for Teacher Dependent Variables Used in Multilevel Models, by State, 2005–2006.

	California			Georgia			Pennsylvania		
	Teacher	School	District	Teacher	School	District	Teacher	School	District
Test-focused instructional practices	.95	.00	.05	.95	.00	.05	.87	.11	.02
Use of achievement data	.96	.04	.00	.93	.04	.02	.84	.13	.02

consistent with the item-level results shown in Table 9, the bulk of the variance is observed among teachers at the same school rather than between schools or districts, leaving relatively little variance to explain at higher levels.

Although the variance decomposition results suggest an absence of strong school- or district-level effects on teachers' practices, it is still possible to construct multilevel models that attempt to explain the relatively small proportion of variance at either the school or district level. Moreover, the reliability of the teacher practice measures increases as a result of the decision to aggregate the responses of teachers within each school. Tables 11 and 13 show the results of the models predicting each teacher practice measure from school-level variables. As discussed earlier, decisions about which principal-reported measures to include in each model were based on an initial set of exploratory models. The tables provide the standardized coefficients, standard errors, *p*-values, and 95% confidence bands for each coefficient.

Table 11 presents the results for *test-focused instructional practices*. The coefficient for the California indicator variable was negative and significant, which suggests that on average, teachers in California reported less of an influence of testing on their practices than did teachers in Pennsylvania (the omitted category) after controlling for the other variables in the model. The coefficient for the indicator variable for Georgia was not significant. The significant coefficient for elementary schools implies that there was more test-focused instruction in elementary schools than middle schools.

In addition, principal-reported AYP pressure is statistically significant, indicating that there was more test-focused instruction in schools where principals felt strong AYP pressure than in schools where principals did not feel strong AYP pressure. Additionally, there was less test-focused

**Table 11.** Fixed Effect Coefficients for Multilevel Model Predicting Test-Focused Instructional Practice.

Predictor	Estimate	Standard Error	<i>p</i>	95% Confidence Interval	
State = CA	-0.17	0.06	0.01	-0.29	-0.05
State = GA	0.03	0.06	0.58	-0.08	0.15
Year	0.09	0.02	<.0001	0.05	0.13
Elementary school	0.16	0.03	<.0001	0.10	0.23
Principal: AYP pressure	0.03	0.01	0.01	0.01	0.06
Principal: Instructional leadership	0.03	0.02	0.08	0.00	0.07
Principal: Use achievement data	-0.04	0.02	0.11	-0.09	0.01
Principal: Use research	0.03	0.03	0.23	-0.02	0.09
Made AYP	-0.11	0.03	0.00	-0.18	-0.05
Constant	-0.18	0.09	0.04	-0.35	-0.01

**Table 12.** Random Effect Coefficients for Multilevel Model Predicting Test-Focused Instructional Practice.

Random Effect	Standard Deviation
Between school intercept	0.02
Between school slope	0.13
District intercept	0.16
Residual	0.19

instruction in schools that made AYP. The positive coefficient for “year” indicates that test-focused instruction increased between 2004–2005 and 2005–2006, controlling for the other variables in the model. Table 12 shows the random (or varying) coefficient. The average effect of year, as shown in Table 11, is 0.09, which means that the “average” school increased by 0.09 (controlling for other parameters). The standard deviation of this estimated average effect is 0.13, as shown in Table 12. We would expect approximately two-thirds of schools to lie within one standard deviation of the mean effect, which means that the lowest 1/6th of schools have a slope at or above  $0.09 - 0.13 = -0.04$ , and the upper 1/6th have a slope at or above  $0.09 + 0.13 = 0.22$ . Therefore there is considerable variation in the rate of increase, with some schools decreasing over time and others increasing to varying degrees.<sup>7</sup> Neither of the school improvement strategies included in the model – use of achievement data to improve instruction and use of existing research to inform decisions about improvement – was significantly

associated with test-focused instruction. The coefficient for instructional leadership also failed to reach significance.

Tables 13 and 14 present the results for teacher’s *use of achievement data*. The coefficients for both state indicator variables were significant, which suggests that on average, teachers in California and Georgia were more likely to report using assessment data to guide instruction than were teachers in Pennsylvania (the omitted category) after controlling for the other variables in the model. The results also indicate that use of achievement data increased over time, controlling for the other variables in the model. In addition, teachers in elementary schools were more likely to report using achievement data than were those in middle schools. Teachers in schools where principals reported increasing the quantity of professional development reported less data use. Because we lack information on the nature of the professional development, it is difficult to interpret this relationship; it is possible, for example, that the professional development

**Table 13.** Fixed Effect Coefficients for Multilevel Model Predicting Use of Achievement Data.

Predictor	Estimate	Standard Error	<i>p</i>	95% Confidence Interval	
State = CA	0.12	0.06	0.035	0.01	0.23
State = GA	0.25	0.05	<.0001	0.14	0.35
Year	0.13	0.02	<.0001	0.09	0.18
Elementary school	0.23	0.04	<.0001	0.16	0.30
Principal: Support for data use	0.02	0.03	0.435	-0.03	0.07
Principal: Increase quantity of professional development	-0.08	0.03	0.022	-0.14	-0.02
Constant	-0.38	0.08	<.0001	-0.55	-0.22

**Table 14.** Random Effect Coefficients for Multilevel Model Predicting Use of Achievement Data.

Random Effect	Standard Deviation
Between school intercept	0.15
Between school slope	0.08
District intercept	0.12
Residual	0.24

most teachers experienced emphasized topics other than data use. These findings will be explored further in future analysis of survey and case study data.

These modeling results, while limited to a small number of predictors and outcomes, illustrate some of the ways in which our multilevel data permit explorations of relationships among responses to SBA across levels of the education system. While the findings provide further evidence that NCLB has overcome some barriers to external coordination of teachers' instruction, they also suggest the persistence of some loose coupling between classrooms and school administration. The next step in this analysis will involve refining the models above. Cross-lagged regression models using all three waves of data will provide a more powerful means to examine the relationships between school factors and teacher practice. In addition, we will estimate models with student achievement as a dependent variable, and these models will include predictors at the district level.

## DISCUSSION

These findings suggest that teachers' practices and attitudes are being shaped in multiple ways by their state SBA systems. Much of the influence appears to be direct: teachers are responding individually to pressures from the accountability system. Consistent with research on school organizations, little evidence suggests that specific district and school actions are affecting teachers' own responses. The large amount of within-school variance in most of the practice and attitude measures indicates that teachers continue to exercise a large degree of autonomy over their instruction, even in districts that have adopted uniform policies around curriculum, instruction, and professional development. However, evidence of an influence on some teachers' instructional practices suggests NCLB has overcome some barriers to the limited governability of teaching.

The multilevel modeling described here was preliminary, but suggests that the effects of principal attitudes and actions on teachers' practices are generally weak, as they must be in the presence of high within-school variability. The relationships that were observed tend to be consistent with widely held views about how school-level conditions are likely to affect what teachers do. For example, principals' sense of AYP-induced pressure was associated with teacher reports of practices focused on tests, as was a school's failure to make AYP the previous year. However, we cannot tell from these data whether principals' sense of pressure directly

influenced teachers' actions, or whether other conditions (such as persistent low test scores) affected both principal and teacher responses. Additional analyses that incorporate information about achievement and that use multiple waves of data will illuminate the nature of this relationship.

These analyses reflect other limitations, including the lack of a design that supports causal inference. Additional analyses exploiting the longitudinal nature of the study could provide somewhat stronger evidence regarding causality but will not solve the fundamental design limitations. The absence of relationships observed in some instances may stem in part from weaknesses in our measures of practices and school improvement strategies, which do not capture variation in the quality of these actions. Some of the findings could be illuminated through subsequent small-scale studies that use richer (but more expensive) practice measures. The study is further limited by the inclusion of only three states and a relatively small (though representative) sample of districts in those states.

SBA policies aim to make classroom practices more governable by central actors by defining and measuring outcomes discretely, and then using the resulting data to shape instruction. Our findings, along with those of other recent studies of teachers' responses to SBA, suggest some shift away from the ungovernability of teaching described by sociologists (Dreeben, 1970; Firestone, 2003). SBA policies enacted at the federal and state levels are clearly being felt in the classroom and are influencing what teachers do on a day-to-day basis. Although teachers maintain a good deal of autonomy, it is constrained: they report feeling pressure to make certain changes to their instruction, even when they view these changes as inconsistent with their personal views of how they should teach.

However, school and district policies continue to have weak effects on teachers' classroom practices. Teachers respond to SBA, but they do in varying ways and not in harmony with their colleagues. It may be the case that school-level influences will grow stronger under NCLB as annual performance targets increase and district interventions intensify. As schools become identified for improvement, the nature of the interventions becomes more centralized. With corrective action and restructuring, the role of the school and district grows even stronger. One might expect to see more powerful district and school actions in less successful schools, and one might also expect to see shifts toward tighter coupling of school systems as NCLB policies continue to be implemented.

Our findings contribute to a growing number of studies demonstrating how the technical environment of schooling – seriously altered by SBA



policies – penetrates classroom boundaries. Yet, the findings also show that policy mandates continue to have an uncertain relationship with teacher practice. Although the specific reasons for this are unknown, there are several potential explanations in the context of SBA. For example, while state standards and assessments provide some guidance for teacher action, they still lack procedural templates of sufficient specificity to guide teachers' day-to-day practice. In addition, varying responses to policy are likely to some extent to be the result of the cognitive demands of this ambitious reform policy. Teachers must first make sense of what policy asks (Coburn, 2001; Spillane, 2004). In addition, changes in practice require teachers to learn new skills and behaviors, and such learning requires time and is highly individualistic. Further, NCLB does little to directly alter the loosely coupled structure of educational organizations, and school and district administration has a relatively weak impact on classroom instruction (Bidwell, 2001). As we noted earlier, some SBA policies have reduced teachers' autonomy and control over their classrooms, but they have not shifted this control to principals or districts. A further mismatch occurs because NCLB targets schools as the unit of intervention, but requires changes in teacher behavior at the classroom level.

Much more remains to be learned about the relationship between teachers' classroom practice and policies emanating from school administrators and higher levels of the educational hierarchy. Given that prior research suggests schools have the potential to shape teachers' responses to policy in positive directions, it is critical that we gain further insight into the mechanisms that promote schoolwide instructional improvement efforts in the context of NCLB and other accountability policies.

## NOTES

1. Instructional leaders are knowledgeable about instruction and therefore able to lead, directly and indirectly support, and hold teachers accountable for implementation of standards, curriculum reforms, and other instructional improvement initiatives (Spillane, Halverson, & Diamond, 1999; Supovitz & Poglinco, 2001).

2. Similar findings from earlier years can be found in Hamilton et al. (2007).

3. Results for science teachers will be presented in later project reports.

4. There is also a case study component to the project. We interviewed principals, teachers, and parents at selected schools in each state. The case study results are reported elsewhere (Hamilton et al., 2007).

5. Because our surveys were customized to each state, we were able to include the actual name of the state testing program; e.g., the Pennsylvania System of School Assessment or PSSA.

6. Measurement error is also included in this component.

7. This is making the assumption that these effects are approximately normally distributed. All such assumptions were tested and were found to be satisfied to a reasonable degree.

## REFERENCES

- Abelmann, C., Elmore, R., Even, J., Kenyon, S., & Marshall, J. (1999). *When accountability knocks, will anyone answer?* No. CPRE Research Report Series RR-42. Consortium for Policy Research in Education.
- Bidwell, C. E. (2001). Analyzing schools as organizations: Long-term permanence and short-term change. *Sociology of Education*, 74, 0, Extra issue: Current of thought: Sociology of education at the dawn of the 21st century, 100–114.
- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas accountability system. *American Educational Research Journal*, 42, 231–268.
- Center on Education Policy. (2006). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Centre on Education Policy. Available at <http://www.cep-dc.org/nclb/Year4/Press/>
- Coburn, C. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170.
- Cohen, D. K. (1996). Rewarding teachers for student performance. In: S. H. Fuhrman & J. A. O’Day (Eds), *Rewards and reform: Creating educational incentives that work* (pp. 60–112). San Francisco: Jossey-Bass.
- Corbett, H. D., & Wilson, B. L. (1991). Two state minimum competency testing programs and their effects on curriculum and instruction. In: R. E. Stake (Ed.), *Advances in program evaluation: Vol. 1. Effects of mandated assessment on teaching* (pp. 7–40). Greenwich, CT: JAI Press.
- Darling-Hammond, L., & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *Elementary School Journal*, 85, 315–336.
- Dobbs, M. (2004). ‘No child’ law leaves schools’ old ways behind. *The Washington Post*, April 22, p. A1.
- Dreeben, R. (1970). *The nature of teaching: Schools and the work of teachers*. Glenview, IL: Scott, Foresman.
- Firestone, W. A. (1985). The study of loose coupling: Problems, progress and prospects. In: A. Kerckhoff (Ed.), *Research in sociology of education and socialization* (Vol. 5, pp. 3–30). Greenwich, CT: JAI Press.
- Firestone, W. A. (2003). The governance of teaching and standards-based reform from the 1970s to the new millennium. In: M. T. Hallinan, A. Gamoran, W. Kubitschek & T. Loveless (Eds), *Stability and change in American education: Structure, process, and outcomes* (pp. 153–170). Clinton Corners, NY: Eliot Werner Publications.
- Firestone, W. A., Monfils, L. F., Hayes, M., Polovsky, T., Martinez, M. C., & Hicks, J. E. (2004). The principal, test preparation, and educational reform. In: W. A. Firestone, L. F. Monfils & R. Y. Schorr (Eds), *The ambiguity of teaching to the test: Standards, assessment, and educational reform* (pp. 91–112). Mahway, NJ: Lawrence Erlbaum Associates.

- Goodlad, J. I. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.
- Hamilton, L. S. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25–68.
- Hamilton, L. S., Stecher, B. M., Marsh, J., McCombs, J. S., Robyn, A., Russell, J., Naftel, S., & Barney, H. (2007). *Implementing standards-based accountability under No Child Left Behind: Responses of superintendents, principals, and teachers in three states*. Santa Monica, CA: RAND.
- Hill, H. C. (2005). Content across communities: Validating measures of elementary mathematics instruction. *Educational Policy*, 19(3), 447–475.
- Ingersoll, R. M. (2003). *Who controls teachers' work? Power and accountability in America's schools*. Cambridge, MA: Harvard University Press.
- Jones, G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81, 199–203.
- Kirby, S. N., Berends, M., & Naftel, S. (2001). *Implementation in new American schools: Four years into scale-up*. Santa Monica, CA: RAND.
- Knapp, M. S., Copland, M. A., & Talbert, J. E. (2003). *Leading for learning; reflective tools for school and district leaders*. Seattle, WA: Center for the Study of Teaching and Policy.
- Koretz, D., Barron, S., Mitchell, K., & Stecher, B. (1996). *The perceived effects of the Kentucky instructional results information system (KIRIS) (MR-792-PCT/FF)*. Santa Monica, CA: RAND.
- Koretz, D. M., & Hamilton, L. S. (2003). *Teachers' responses to high-stakes testing and the validity of gains: A pilot study*. CSE Technical Report 610. Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
- Le, V., Stecher, B. M., Lockwood, J. R., Hamilton, L. S., Robyn, A., Williams, V., Ryan, G., Kerr, K., Martinez, F., & Klein, S. (2006). *Improving mathematics and science education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement*. Santa Monica, CA: RAND.
- Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. Minneapolis, MN: Center for Applied Research and Educational Improvement, University of Minnesota.
- Lemons, R., Luschei, T. F., & Siskin, L. S. (2003). Leadership and the demands of standards-based accountability. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *The new accountability: High schools and high-stakes testing* (pp. 87–127). New York: Routledge Falmer.
- Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509–536.
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.
- Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers and mathematics and science: The state of the scene. *School Science and Mathematics*, 99(5), 258–271.
- Marks, H., & Louis, K. S. (1997). Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educational Evaluation and Policy Analysis*, 19, 245–275.
- Mayer, D. (1999). Measuring instructional practice: Can policymakers trust survey data? *Educational Evaluation and Policy Analysis*, 21, 29–45.
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83, 340–363.

- Meyer, J. W., & Rowan, B. (1978). The structure of educational organizations. In: M. W. Meyer (Ed.), *Environments and Organizations* (pp. 78–109). San Francisco: Jossey-Bass Publishers.
- Mullens, J. E., & Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction* (NCES, 1999-08). Washington, DC: National Center for Education Statistics.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from a national survey of teachers*. Boston: National Board on Educational Testing and Public Policy.
- Romberg, T. A., Zarinia, E. A., & Williams, S. R. (1989). *The influence of mandated testing on mathematics instruction: Grade 8 teachers' perceptions*. Madison, WI: National Center for Research in Mathematical Science Education, University of Wisconsin – Madison.
- Rothman, R., Slatery, J. B., Vranek, J. L., & Resnick, L. B. (2002). *Benchmarking and alignment of standards and testing*. CSE Technical Report No. 566. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA.
- Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. In: C. Cazden (Ed.), *Review of research in education*, 16. Washington, DC: American Educational Research Association.
- Rowan, B. (1995). Research on learning and teaching in K-12 schools: Implications for the field of educational administration. *Educational Administration Quarterly*, 31, 115–133.
- Rowan, B., & Miskel, C. (1999). Institutional theory and the study of educational organizations. In: J. Murphy & K. S. Louis (Eds), *Handbook of research on educational administration* (2nd ed., pp. 359–383). San Francisco: Jossey-Bass Publishers.
- Shepard, L. A., & Dougherty, K. C. (1991). Effects of high-stakes testing on instruction. Paper presented at the annual meeting of the American Educational Research Association and National Council on Measurement in Education, Chicago.
- Smith, M. L., Edelsky, C., Draper, K., Rottenberg, C., & Cherland, M. (1991). *The role of testing in elementary schools*. CSE Technical Report 321. Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In: S. H. Fuhrman & B. Malen (Eds), *The politics of curriculum and testing (Politics of Education Association Yearbook, 1990)* (pp. 233–267). London: Taylor and Francis.
- Smithson, J. L., & Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum – the reform up close Study*. CPRE Research Report Series #31. Consortium for Policy Research in Education, University of Wisconsin, Madison, WI.
- Spillane, J. P. (2004). *Standard deviation: How schools misunderstand education policy*. Cambridge, MA: Harvard University Press.
- Spillane, J. P., Halverson, R., & Diamond, J. (1999). Distributed leadership: Toward a theory of school leadership practice. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.
- Stecher, B. M., Barron, S. I., Chun, T., & Ross, K. (2000). *The effects of the Washington state education reform on schools and classrooms*. CSE Technical Report 525. Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.

- Stecher, B. M., Hamilton, L. S., & Gonzelez, G. (2003). *Working smarter to leave no child behind*. Santa Monica, CA: RAND.
- Stecher, B. M., Hamilton, L. S., & Naftel, S. (2005). *Introduction to first-year findings from the implementing standards-based accountability (ISBA) project*. Santa Monica, CA: RAND.
- Supovitz, J. A., & Poglianco, S. M. (2001). *Instructional leadership in a standards-based reform*. Philadelphia, PA: Consortium for Policy Research in Education.
- Waters, T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement*. Aurora, CO: Mid-continent Research for Education and Learning.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1–19.

## CHAPTER 3

# DISTRICT LEADERS ERODING SCHOOL COHERENCE? THE INTERPRETATION OF ACCOUNTABILITY MANDATES

Thomas F. Luschei and Gayle S. Christensen

### ABSTRACT

*We examine how school districts in California help their high schools respond to state accountability requirements. We discovered two contrasting forms of district interventions: those aiming to increase schools' internal coherence and those encouraging direct but narrower responses to state requirements. Drawing on interviews in six districts and eight high schools, we find that many district efforts focus on immediate responses to state requirements to raise test scores. Yet, our analysis suggests that without strong district efforts to increase internal coherence, interventions aimed at eliciting school responses will be less beneficial over time.*

### INTRODUCTION

A major aim of accountability policies such as the No Child Left Behind (NCLB) Act and state accountability systems is to elicit positive

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
**Research in Sociology of Education, Volume 16, 67–101**  
Copyright © 2008 by Emerald Group Publishing Limited  
All rights of reproduction in any form reserved  
ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16003-7

responses – through both sanctions and rewards – from schools falling below externally determined criteria. Yet, growing evidence indicates that schools have responded to such policies in unpredictable and varied ways (Abelmann & Elmore, 1999; DeBray, Parson, & Avila, 2003). While some schools respond to increased pressures from federal and state governments by mobilizing campaigns to prepare for tests, others have sought deeper changes, such as greater alignment of the curriculum with standards, analysis of student test score data to inform instruction, and the provision of more challenging curriculum to a wider population of students. Still other schools demonstrate almost no visible response to growing pressures to raise test scores.

Previous research on accountability in the United States has offered several explanations for variation in how accountability policies are interpreted and enacted by local educators. These include differences in schools' capacity to respond with effective and relevant solutions, schools' internal coherence, the strength of stakes brought to bear by the state, and the immediacy of sanctions faced by schools falling below externally determined criteria (Abelmann & Elmore, 1999; DeBray et al., 2003; Elmore, 2003; Luschei, Goldwasser, Carnoy, Christensen, & Weinbaum, 2007; Massell, Goertz, Christensen, & Goldwasser, 2005). Organizational theorists have also explored the degree to which schools are "loosely coupled" organizations, in which decisions are not closely tied to implementation and individual elements maintain considerable independence from each other (Weick, 1976, 1982). Although few researchers have applied the idea of loose coupling to schools' responses to accountability policies, there is clearly a conceptual link between the tightness of coupling within an educational system and the responsiveness of the system to external mandates or calls for change (Gamoran & Dreeben, 1986; O'Day, 2002; Orton & Weick, 1990).

While previous explanations shed light on variations in school response, they often neglect the role of critical actors in the state–school relationship: leaders situated in school district offices. Our research has found that districts often mediate accountability policies and pressures as they make their way from the state to the school. Yet, few researchers have examined the extent to which districts facilitate, encourage, or demand that schools respond to the state's requirements. In a large state with a relatively high-stakes accountability system like California, the district can play many different and important roles vis-à-vis the school's response to state requirements. For example, a district can offer schools instructional resources, professional development for teachers, or assistance in the collection and analysis of student test score data. Districts have also played

a role in selecting and defining the roles of external actors who work directly with low-performing schools.

California has a wide variety of district types and sizes, including K-8 districts, high-school-only districts, and unified districts. Each type is likely to have its own approach and resources to draw on as it works with its schools to meet state accountability requirements. In this study, we make use of a diverse sample of eight California high schools and six districts – including two high-school-only districts and two districts with two schools in our sample – to explore (1) whether districts are responding to state accountability policies, (2) whether districts direct their efforts more toward improving schools’ internal coherence or eliciting external responses, (3) how districts’ responses vary according to their own characteristics and those of their schools, and (4) which district responses, if any, influence schools to respond in ways that the state expects. We also explore the degree to which our data and analysis support the ideas of Weick (1976, 1982) and others that schools are loosely coupled institutions that cannot be easily influenced by external mandates from the district or state. We conclude with a discussion of the relevance of organizational theory perspectives in shaping districts’ and states’ directives to encourage schools to respond and improve.

## **THEORETICAL FRAMEWORK**

### *The Role of the District*

The expressed purpose of state accountability systems is to raise student achievement and, more generally, to improve the quality of schooling. Some accountability systems, including NCLB, also hold districts accountable for the performance of their schools. Not surprisingly, researchers have debated the importance and effectiveness of school districts in this effort (and in general). Chubb and Moe (1990) argue that districts have strong incentives to use bureaucratic controls to encourage school compliance, a “clumsy and ineffective” way of delivering education. Other researchers have shown that districts can play a crucial role in state reform efforts (Fuhrman & Elmore, 1990). Yet, regardless of their organizational effectiveness, districts are written into accountability policies as active players with specific requirements, and many accountability systems implicitly or explicitly increase the role that they play in providing assistance and monitoring performance (Weinbaum, 2005). District activities with respect to accountability requirements have included emphasizing particular aspects of policies and building



structures to implement requirements at the local level (Goertz, Massell, & Chun, 1998; Marsh et al., 2005).

Recent research indicates that although districts often respond to state accountability policies, there are large variations in their responses (Marsh et al., 2005; Weinbaum, 2005). These studies find a range of district activities aimed at improving schools' performance, including efforts at improving curriculum and instruction, professional development, and greater use of data. These studies also find differences in response based on a range of district and school contextual factors, such as demographic characteristics and school performance. In general, however, these studies have not linked district actions and characteristics to response in high schools, especially in districts with schools performing below the state average.<sup>1</sup> One exception is a survey study of schools in California that found that schools' improvement on state accountability measures varied by district characteristics (Adamson, Carnoy, Addy, Ricalde, & Rhodes, 2007). This study found that schools in unified school districts had significantly lower improvement in their state accountability scores than schools in other types of districts, such as high-school-only districts. A separate case study of three diverse California districts conducted by Policy Analysis for California Education offers three lessons learned for state and district policymakers:

- (1) Coherence and engagement across all levels of the district promote sustainability of reforms.
- (2) Districts' limited resources must be focused on their reform goals.
- (3) Low-performing districts serve very diverse student populations; reform efforts must address gaps in achievement across race and socioeconomic status (Woody, Bae, Park, & Russell, 2006, pp. 8–9).

The report's emphasis on coherence underscores the importance of districts' organizational and cultural features. We argue that districts must also consider coherence within the *school*. Given the new accountability systems, there is a clear need for districts to help schools take actions to meet new state standards. But this focus may not be sufficient to help schools make substantial strides in their accountability performance. Work conducted by researchers from the Consortium for Policy Research in Education (CPRE) has found that schools with weaker organizational coherence or *internal accountability* are less likely to respond to state mandates (Carnoy, Elmore, & Siskin, 2003; Luschei et al., 2007). According to Abelmann and Elmore (1999), schools have conceptions of accountability embedded in the patterns of their day-to-day operations that significantly influence how they deliver education. In addition to the pressures of external

accountability systems – such as state or district bureaucracies or market forces – many school stakeholders influence a school’s internal accountability.

As Elmore (2003) argues, the nature and magnitude of a school’s internal accountability is likely to influence its response to external requirements:

...all schools have deep-seated norms and predispositions that determine their conceptions of accountability. It is not the case that some schools are accountable and others are not. All schools are accountable, but different schools solve the accountability problem in very different ways. Many schools have very diffuse notions of being accountable “to the children,” which often ends up meaning that individual teachers enact their own views of what their students need, unmediated by collective views within the school about what the organization believes, or what parents demand, or what public policy requires. A few schools have strong collective views of what they stand for, and well-developed organizational processes that bring those beliefs into action. We characterize the former schools as having weak, and the latter as having strong, *internal accountability*. It became clear to us that the strength and focus of internal accountability in schools was a key determinant in how they would respond to any external accountability system. (p. 196)

### *Internal Accountability and Loose Coupling*

From the lens of organizational theory, internal accountability can be seen as an indicator of the degree of coupling around a shared view of the mission or direction of the school. The idea of varying degrees of coupling arose from the reactions of Weick (1976) and others to the application of a Weberian rational bureaucratic model of “coherence, coordination, and control” to education (see description in Meyer, 1986). In contrast to the rational bureaucratic model, these theorists argued that schools are characterized by loose coupling among events and actors (March & Olsen, 1976; Weick, 1976, 1982). Elaborating on the concept of loose coupling, Weick (1982) explains that, in loosely coupled systems, “ties among people are weaker, more unpredictable, and more intermittent than is true of other forms of organization in which directives diffuse more rapidly and deviations are dealt with more swiftly” (p. 676). Weick also argues that, although certain aspects of schools are in fact tightly coupled, such as the bus schedule and the payment of personnel, “the task of educating is simply not the kind of task that can be performed in a tightly coupled system” (1982, p. 674). School administrators must recognize this looseness of coupling and concentrate on building tight coupling around key values. Of particular importance in this task are the administrator’s “voice and vision”:

Diverse ideas about the school’s mission are common under conditions of loose coupling. This very plurality makes for successful local accommodations. But people

also need some shared sense of direction for their efforts. This already exists in most tightly coupled systems, but it must be built and reaffirmed in a loosely coupled system. Articulating a theme, reminding people of the theme, and helping people to apply the theme to interpret their work – all are major tasks of administrators in loosely coupled systems. (Weick, 1982, p. 675)

Our conception of internal accountability draws on Weick's idea regarding the work of administrators and other school actors to develop a shared sense of direction, albeit in a loosely coupled system. This shared sense of direction in many ways parallels Elmore's (2003) idea of internal accountability. Recognizing that the schools we observed in California feature a high degree of loose coupling, we are particularly interested in the extent to which schools are able to develop and organize around a common sense of direction. We believe that this ability is not only important for the success of school administrators, but also a key determinant of a school's response to external mandates.

#### *Loose Coupling and Policy Responses*

Assumptions that accountability policies will elicit parallel and rapid responses in schools rest on the notion that external policy and internal action are tightly coupled. Yet, as discussed above, this Weberian notion of bureaucratic coherence has been challenged by researchers who argue that schools and school systems follow a much different model, characterized by loose coupling among organizational subunits (March & Olsen, 1976; Weick, 1976, 1982). Competing visions of the degree of coupling of educational systems clearly have important consequences for the success of policies that assume rational and immediate response. For example, researchers have found links between loose coupling and greater persistence, manifested by resistance to change and reduced responsiveness of organizations, while tightly coupled systems are more apt to respond to systemwide change (Firestone, 1985; Orton & Weick, 1990; Wilson & Corbett, 1983). Other researchers have found that educational organizations differ in the tightness of coupling and the functions around which they are coupled (Gamoran & Dreeben, 1986). Meyer, Scott, and Deal (1983) argue that while educational organizations are generally loosely coupled around functions pertaining to the coordination of work, they are more tightly coupled around institutional rules. In an empirical examination of this dichotomy, Gamoran and Dreeben (1986) conclude that while tightness of

connections varies across districts, district and school administrators can exercise indirect control over the technical instructional core within classrooms through the allocation of resources. Their study demonstrates that while making changes to the instructional core of teaching presents a strong challenge to administrators, it is nonetheless possible.

In the context of recent accountability reforms in the United States, researchers have argued that loose coupling within the educational system impedes the success of external mandates (McDermott, 2006; Smith & O'Day, 1990). O'Day (2002) observes that an inherent weakness of outcomes-based school accountability policies in the United States is that they “operate from a bureaucratic control model and thus fail to create the interaction patterns and normative structures within schools that encourage sustained learning and adaptation” (p. 315).

These arguments point out a key tension in recent accountability reforms: while schools are generally loosely coupled and difficult to manage either internally or from above, accountability reforms assume more tightly coupled arrangements within and across levels.

Our experience in California suggests that schools are indeed loosely coupled systems in which subunits such as departments and individual teachers are independent and occasionally in opposition with each other. At the same time, many district efforts to encourage response assume tightly coupled environments. In effect, the two levels operate under conflicting models. Not surprisingly, the result is often misunderstanding between districts and schools. As Weick (1982) argued, school stakeholders “apply the only model they know – the rational bureaucratic model – and when they don’t see the school measure up to these standards, they look for someone to blame or someone to fire” (p. 674). Similar to Weick, we argue that in order to effect instructional changes, the district must recognize the importance of looseness of ties among school actors. This is particularly important in effecting change within the technical instructional core of schools, or the actual classroom work and practices of teachers.

Traditionally, changes to the instructional core have been regarded as very difficult to make, and as organizational theorists have observed, administrators’ work is often not strongly connected to this technical core (e.g., Meyer, 1980; Meyer & Rowan, 1978). Yet, Gamoran and Dreeben (1986) argue that, because the technical core of what teachers teach and what students learn has rarely been documented, it is not clear that administrators do not exert any influence on classroom instruction. In their study of three Chicago school districts, Gamoran and Dreeben (1986) found

that, through the manipulation of resources, district and school administrators do indeed affect teachers' work in the classroom. However, due to loose coupling within educational systems, "technological connections do not appear as direct controls from the central office to the classroom" (p. 5). Instead, school actors mediate the effects of district decisions on classroom instruction.

Although Gamoran and Dreeben (1986) found that districts can influence classroom instruction through the allocation of resources, there is little evidence regarding the influence of more symbolic efforts. In particular, can the development of a common vision around instruction, as described by Weick (1982), influence instructional changes? Can districts aid in the development of this vision? In our study, we use the concept of internal accountability as an indicator of a common vision around instructional change. Additionally, we conceptualize district activities to influence internal accountability as efforts to change the instructional core indirectly – most likely, through support to school leaders. In contrast to this more subtle and indirect approach, our research indicates that district actors more commonly attempt to elicit immediate responses to accountability policies, such as increased test preparation. While the second type of effort is more direct, the first better addresses the organizational looseness of most schools and thus may be more effective.

Drawing on data from six districts and eight schools, we examine the influence of district efforts on school responses to accountability policies. By using data from one state, we are able to focus on efforts within a single-state accountability context. Of course, even within a similar policy framework, districts and schools may interpret and experience state policies differently. Despite this interpretive variability, the same terms and requirements apply to all schools in the study. Our conceptual framework is based on current research on districts' and schools' responses to accountability systems (Fig. 1). We posit that the district may provide supports in two areas: actions to increase internal accountability and actions to encourage external response.

We also believe that schools may influence district responses by asking districts to provide particular supports to improve either internal accountability or external response. Efforts at the district or school level may be influenced by both the district and school context. For example, a district may choose to focus more attention on schools that are designated by the accountability system as in need of improvement. Further, a high-school-only district may be able to provide more effective supports to high

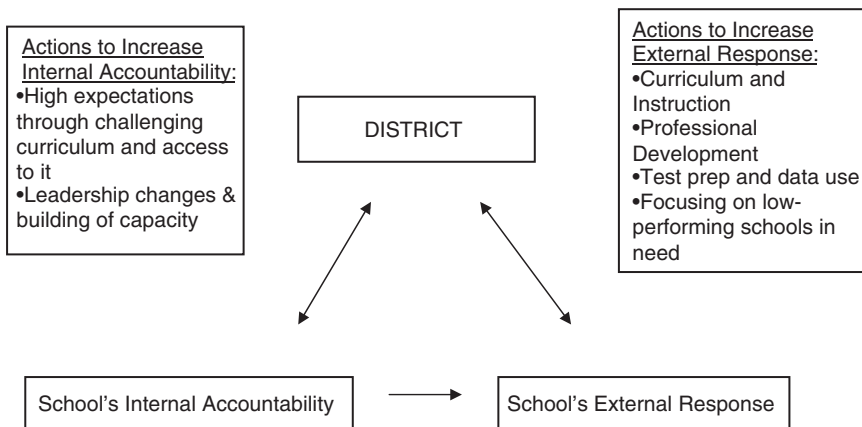


Fig. 1. Conceptual Framework.

schools due to greater knowledge and experience of these complex organizations. At the school level, a high school can be in a district with many activities aimed at external response, but if the school has weak internal accountability, it may be unable to make effective use of this support. Overall, our framework suggests that in order for district efforts to be successful, district leaders must take a two-pronged approach to improve both a school's internal accountability and its external response.

## RESEARCH QUESTIONS

Based on the background and conceptual framework outlined above, this study seeks to answer three research questions regarding how districts respond to state accountability systems and how district and school responses interact:

- (1) Are districts responding to state accountability policies and, if so, how do responses vary according to district characteristics?
- (2) How do responses vary according to *school* characteristics? For example, do districts provide different supports to schools that are in need of improvement based on accountability measures?
- (3) What is the relationship between district efforts and schools' responses?

## DATA AND METHODS

### *Sampling Strategy and Data*

The analysis of California schools described here is based on a larger cross-state effort by CPRE to examine how high schools respond to accountability policies across a wide variety of contexts. In order to capture variability in both schools and states, the research team selected eight schools in each of six states: California, Florida, Michigan, North Carolina, New York, and Pennsylvania. At the time of data collection, two of these states (Pennsylvania and Michigan) were considered “weak” accountability states, while the remaining four were “strong” accountability states.<sup>2</sup> Within each state, eight schools were sampled based on mathematics test score data from the 1999–2000 school year and a number of context indicators. In California, context indicators included pupil–teacher ratio, percentage of students receiving free or reduced-price lunches, percentage of minority students, average class size, percentage of teachers with advanced degrees, and average years of teacher experience in a school.

To understand how accountability policies affected low-performing schools that were most likely to feel pressure from the state, the research team selected schools with recent test scores below the state average on mathematics achievement tests. Within this classification, we selected a broad range of schools according to context factors, student performance, and student performance relative to context. Of the eight schools sampled in California, one is rural and the rest are either urban or suburban. These schools represent a wide range of sizes and student compositions. The sample also includes two high-school-only districts, which allows us to examine whether these districts apply greater knowledge or resources to solve the unique problems faced by high schools.

Data collection, which occurred during the 2002–2003 academic year, included visits by multiple researchers to each school in the sample. Researchers conducted structured interviews with school leaders, department leaders in English and Math, and English and Math teachers. To understand the perspectives of those who taught untested subjects, we also interviewed foreign language department chairs in most schools. Interview protocols included topics such as internal and external accountability “press,” external assistance, search for school improvement strategies, and instructional and communal culture. In all six of the districts in our California sample, we also conducted interviews of district officials such as superintendents, assistant superintendents, and chief academic officers.

Coding of interview transcripts centered on key themes like goals and challenges, understanding of accountability policies, responses to accountability, and the nature of improvement strategies used by schools and districts. A common coding scheme across schools and districts facilitated the cross-district analysis.

### *Rating Schools and Districts*

Earlier work relating schools' internal accountability to their responses to state policies has used diverse methods to rate schools in each category. In their study of Kentucky, New York, Texas, and Vermont, DeBray et al. (2003) placed schools on a set of axes: the horizontal axis represented the strength of a school's internal accountability, while the vertical axis indicated alignment with state policy. Along each axis, schools fell along a continuum of low to high, so that, for example, a school with high internal accountability and high alignment with state policy was located in the northeastern quadrant of a diagram relating the two.

Luschei et al. (2007) used a rubric to rate schools numerically based on their internal accountability (*x*-axis) and their alignment with state policy (*y*-axis). Ratings for internal accountability ranged from 0 to 3 for the school's Math department, English department, and the school as a whole. The objective in rating schools was to infer from interview responses and behavior the degree to which actors across the school shared and acted on common goals for students. In rating the English department's internal accountability, for example, the department could receive a 0 ("fragmented/atomized"), a 1 ("bundled"), a 2 ("bonded as a subgroup"), or a 3 ("bonded as a department"). From an organizational theory perspective, these ratings parallel a shared vision or direction. If a subgroup is bonded, for example, this bonding takes place around a vision regarding the school's instructional core. In total, a school could score up to 9 in internal accountability.

Ratings of schools' alignment with state policy were based only on the responses of the Math and English departments. These ratings conformed to the level of actual change to the instructional core that schools implemented. Departments received ratings of 0 ("indifferent/no change"), 1 ("peripheral change"), 2 ("moderate change"), or 3 ("instructional/curricular change"). Schools scored up to a maximum of 6 for alignment with state policy. Using the ratings developed within these rubrics, Luschei et al. (2007) located schools along both the *x*- and *y*-axes as illustrated in Fig. 2.



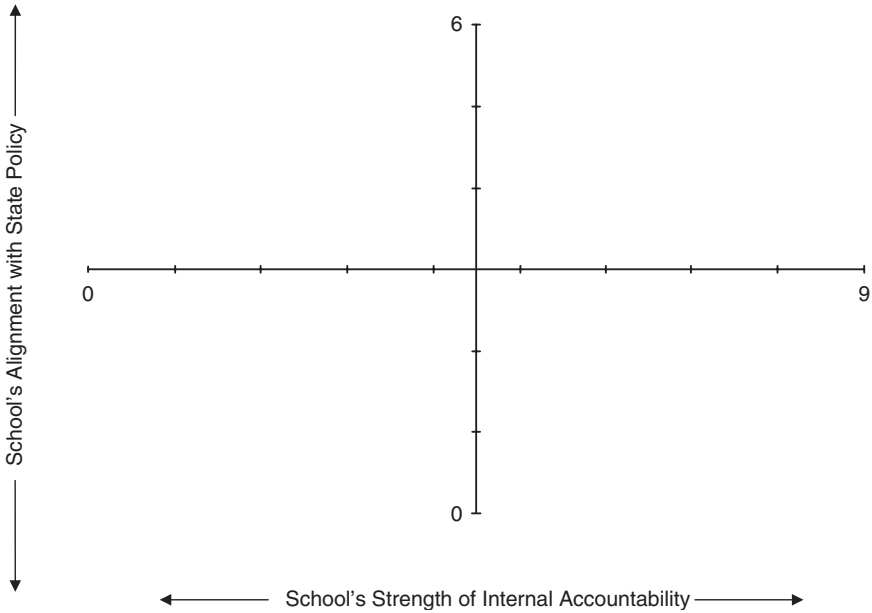


Fig. 2. Internal Accountability and Efforts to Align with State Policies.

To examine the relationship between districts' efforts and schools' responses to accountability, we maintain the California school ratings used by Luschei et al. (2007). We also use a similar rubric to rate the California districts according to their efforts to align their schools with state policies. However, while the earlier school rubrics gave separate ratings for Math and English departments, we rate districts separately based on their efforts directed at low-performing high schools and other high schools in the district. This distinction is based on the work of Weinbaum (2005) and others suggesting that when faced with strong external pressures, districts focus first and most strongly on schools facing the most immediate threats of sanctions. Our ratings allow us to account for differentiated strategies: a district using a strong and comprehensive approach to encourage responses in all of its high schools rates higher than a district with a similar approach directed only at its low-performing schools.

For each type of high school, districts can earn up to 3 points (Table 1). Districts scoring 0 appeared indifferent and exerted little or no effort to increase the response of their high schools to state requirements. Districts with a rating of 1 tried to implement peripheral changes in their schools,

**Table 1.** Rubric for Rating Districts' Efforts to Align Schools with State Policy.

Score	Low-Performing High Schools	Other High Schools
0	Indifferent/no change	Indifferent/no change
1	Peripheral change	Peripheral change
2	Moderate change	Moderate change
3	Instructional/curricular change	Instructional/curricular change

such as after-school tutoring or test preparation activities for students at risk of failing state tests. A district receiving a 2 may have encouraged its schools to make moderate changes, such as lengthening students' courses, adding remedial courses, giving teachers professional development on standards or testing, or making changes to the school schedule. A district receiving the highest rating of 3 focused its strategies on effecting deeper curricular or instructional changes in its high schools, explicitly aligned with state standards and accountability policies. Examples include the adoption of new curricula aligned with state standards, whole-school reform, or the use of student test score data to diagnose problem areas and modify instruction at the individual, class, and school levels.<sup>3</sup>

We do not rate districts' efforts to encourage internal accountability in their schools. Our research questions and interviews with district officials focused on general efforts to help schools respond to accountability. In these interviews, districts primarily demonstrated evidence of efforts to strengthen external responses such as test preparation. Because our data did not yield systematic evidence regarding efforts to improve internal accountability, we cannot adequately rate districts regarding these efforts. However, in the discussion that follows, we mention the few district activities aimed at internal accountability that we did observe.

### *Placing Districts on Axes*

As noted earlier, we use the same ratings and axes as [Luschei et al. \(2007\)](#) to classify the California schools according to internal accountability and response to external policies. To incorporate the actions of districts, we superimpose the district ratings along the vertical axis of alignment to state policy. The resulting diagram allows us to see schools' internal accountability/external response relationship simultaneously with the district's efforts to encourage the school to respond ([Fig. 3](#)). If a district places higher

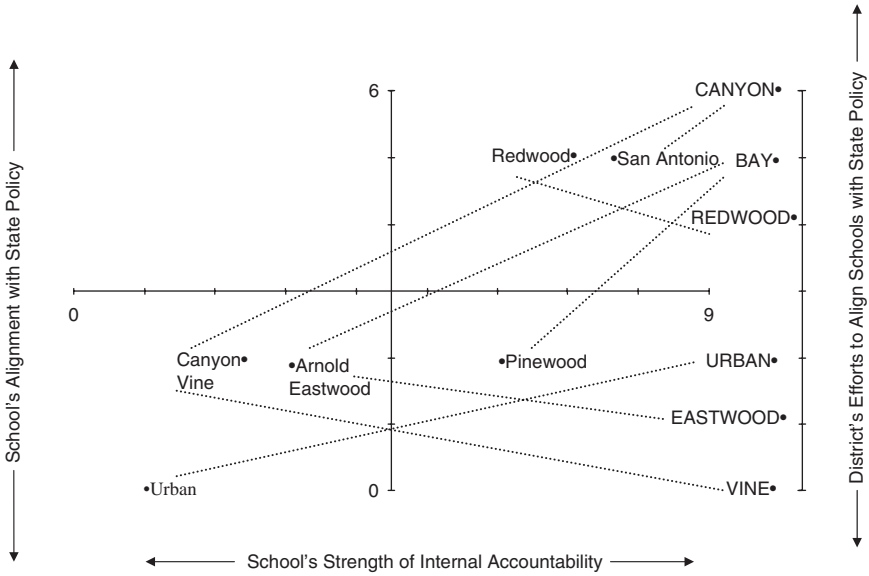


Fig. 3. California Schools and Districts.

on the *y*-axis than its corresponding school(s), this means that the district's efforts fall short of its expectations for schools. In the event that a school places above the district, then the school is somehow exceeding the expectations demonstrated by the district through its own efforts to elicit a response.

## RESULTS

### *State Accountability Context*

California was among four of six states in our original sample with strong accountability systems in 2002–2003. Carnoy and Loeb (2002) gave California a rating of 4 on an accountability strength scale of 1 to 5. Of the 50 U.S. states, only Florida, New Jersey, New York, North Carolina, and Texas scored higher in their rankings. At the time of our interviews in 2002–2003, California's system featured stakes for both schools and students. While schools faced multiple potential sanctions and rewards, the major concern for students was passing the California High School Exit

Exam (CAHSEE). Originally, students in the class of 2004 and beyond were required to pass the CAHSEE in order to receive a high school diploma. However, in July 2003, California's State Board of Education delayed the requirement until 2006. Because we conducted interviews prior to this decision, all of our respondents operated under the 2004 deadline.

California's accountability system in 2002–2003 included both sanctions and rewards for schools, depending on whether they met benchmarks for growth on the Academic Performance Index (API). The API incorporated students' performance on a set of standards- and norm-referenced tests, including the Stanford Achievement Test, 9th Edition (SAT9)/ California Achievement Test, 6th Edition (CAT6) tests. Schools also received a "similar schools" rank comparing their performance to schools with similar student populations. Schools meeting or exceeding their growth targets were originally entitled to monetary rewards, but due to a budget crisis in 2002, most of the incentive programs were discontinued. Repercussions for low-performing schools included identification and voluntary participation in the state's Immediate Intervention/Underperforming Schools Program (II/USP), which provided funds for schools to develop and implement school improvement action plans. II/USP schools failing to demonstrate adequate growth were monitored by the state and were subject to possible sanctions. In 2001 to 2002 academic year, the state replaced II/USP with the High Priority Schools Grant Program (HPSGP) for schools in the lowest decile of API scores.

### *Evidence from Six Districts and Eight Schools*

The six districts in our sample consisted of two high-school-only districts, two districts with two schools each in our sample (including one unified and one high-school-only district), and four districts with only one school in our sample. In the high-school-only district with two sampled schools (Bay District), one had been identified by the state as low-performing and had participated in the II/USP program, while the other was scoring in the middle range of California schools.<sup>4</sup> In the other district with two schools (Canyon District), both schools had been identified as low-performing and were participating in the state's II/USP program. Of the four districts with only one school each, two districts (Urban and Eastwood) had schools that were in "crisis" mode, meaning that issues other than the state's accountability system (such as a state takeover and safety concerns) diverted their attention from state requirements. Both of the two remaining districts (Redwood and Vine) had schools performing at the middle range of California test scores.

*Bay District*

A large high-school-only district, Bay District is located in a rapidly growing urban area. The student population is very diverse ethnically, with large Hispanic and Asian populations. In 2002–2003, about one-quarter of students were English learners and almost one-third qualified for free or reduced-price meals. Bay District contained two of our sample schools, Pinewood and Arnold High. Pinewood, which had a diverse student population distributed among Hispanic, White, and Asian students, was scoring in the middle range of California schools and did not appear to feel acute pressure to raise student test scores. Yet, despite middle-range API scores, the school had a very low similar schools rank, indicating that relative to schools with similar student populations, Pinewood was performing poorly. Arnold High, with a majority of Hispanic students and a large population of English learners, had scored much lower on the API and had been identified as a low-performing school by California's II/USP program. Along with a relatively strong approach on the part of the district oriented toward encouraging school response to state policies, there appeared to be a good working relationship between the district and its high schools.

Bay District had a clear vision of the improvement strategies – focused primarily on getting test scores up – that would work best for its high schools. After identifying a set of target high schools within the district based on recent test score performance, the district provided these schools with a set of professional development and intervention programs to improve the schools' alignment with state policy. As one of the target schools, Arnold received these interventions while Pinewood – due its higher test scores – did not. The district also placed considerable effort on the collection and analysis of data in all of its schools, including the implementation of a fall and spring evaluation of reading, language, and mathematics. This evaluation allowed teachers to track student progress over the course of the year, identifying areas for improvement and making appropriate changes in instruction. The district also revised the curriculum to meet the state standards and instructed all of its high schools to implement the new curriculum. Finally, the district placed a strong emphasis on preparing students for the CAHSEE, including providing resources to give extra tutoring to students who had failed the exam.

Despite Bay District's focus on its low-performing schools, Arnold High was unable to mobilize a strong strategy to meet state requirements. In our interviews, teachers at Arnold tended to express a fatalistic attitude toward

increasing accountability requirements: with the difficult student population they taught, they felt doomed to perform below the state's expectations. As the principal of Arnold High explained, "I feel a lot of heat, not only from the district, from the state and other people ... I guarantee the state can bring in anybody they want in here, they're not going to fix this." Pinewood also had not developed a strong, coherent response to state requirements, despite a similar schools rank indicating that it was underperforming relative to schools with similar student populations. The school relied primarily on efforts to increase student motivation and awareness of tests, rather than substantive organizational or instructional changes. Pinewood also had not undertaken a strong effort to prepare students for the CAHSEE, in contrast to schools in our sample with many students in danger of failing the test. Yet, ironically, despite lower stakes and less attention from the district, Pinewood exerted a less weak response to state requirements than Arnold. For example, while Pinewood teachers made an effort to meet with all of the school's students regarding their test performance, Arnold limited its response primarily to offering extra tutoring for students who had already failed the CAHSEE. These differences may have resulted from growing efforts to encourage inter-departmental collaboration at Pinewood, coupled with the installation of a new leadership team that sought greater cohesiveness across departments. At Arnold, while teachers and administrators paid what one teacher described as "lip service" to improving inter-departmental communication and collaboration, the school seemed fairly atomized.

### *Canyon District*

Canyon District is a large urban unified district. In 2002–2003, a majority of the district's students were classified as English learners and nearly three-quarters qualified for free or reduced-price meals. Canyon contained two of our eight sample schools, San Antonio and Canyon High Schools. At first glance, the two schools were very similar: they were both large comprehensive high schools with large percentages of English learners and students qualifying for free- and reduced-price meals. Both schools had also been identified as II/USP schools and felt a strong accountability press from the state. Yet, while San Antonio had embarked on a coherent, whole-school project to improve instruction and meet state standards, Canyon High was made up of isolated pockets of teachers, occasionally in competition with each other and school leaders. Differences between the schools resulted in

very different relationships with the district: while district officials pointed to San Antonio as an example of the success of their efforts to improve high schools, the same officials expressed concern and frustration at Canyon's inability to coalesce around increased student achievement.

Although Canyon is a unified school district, our interviews with district officials revealed considerable knowledge of high schools, as well as a focused approach to strengthen responses of the district's high schools to external accountability pressures. This approach was guided by a strong commitment to increase students' access to high-quality curriculum, through intensive professional development of teachers in teaching reading through the content areas for its English language learners, as well as expanding offerings of Advanced Placement courses. District officials had also tried to make curriculum more uniform across schools, so that regardless of which school they attended, students graduated with the same body of skills and knowledge. Central in this move toward uniformity was the use of standards-based instruction and assessment. In response to strong accountability pressures from the state – particularly the requirement that all high school seniors pass the CAHSEE to earn a diploma – the district had also developed more immediate responses to raise test scores. These efforts included the implementation of an extra period for freshmen students to prepare for the CAHSEE, providing schools with student test score data to inform instruction, offering after-school tutoring for students who had failed the test, selecting an external evaluator for the schools participating in the state's II/USP and HPSG programs, and developing a guide to help teachers prepare students for the CAHSEE.

While administrators and teachers at San Antonio embraced and complemented the district's efforts, these efforts met with indifference at Canyon. For example, at San Antonio, a strong mathematics department chair had spearheaded efforts to get all freshmen into Algebra 1 and align curriculum to standards. At Canyon, teachers resisted mandatory algebra for freshmen, preferring instead to break Algebra I into a two-year sequence (which lowers a school's accountability score). Differences also manifested themselves in the schools' responses to external requirements. While San Antonio had developed a systematic approach to prepare students for the CAHSEE, including mandatory tutoring for students failing in Math, Canyon's approach was more limited, such as the Math department's plan to spend 10 minutes of each class practicing multiple-choice activities. Additionally, although both schools participated in the II/USP program, teachers at San Antonio considered the program's action plan process to be a valuable experience in identifying areas for improvement and developing

school-wide strategies. At Canyon, teachers deemed II/USP an unnecessary and unhelpful intervention that, in the words of one teacher, “didn’t tell us anything we didn’t already know.” In general, differences in responses between the two schools appeared to correspond to varying degrees of internal accountability. While Canyon’s lack of internal accountability made it difficult for the school to execute an organization-wide response, a strong, shared vision of change at San Antonio facilitated the implementation of a clear instructional improvement plan.

District officials also attributed differences between the two schools to their leadership teams. While San Antonio had a team of leaders with clear roles, Canyon’s relatively new principal appeared to be isolated from other leaders and teachers. There were also large differences in API scores between the two schools: in the year of our interviews, San Antonio jumped from the first (bottom) to third decile in its base score and from the seventh to ninth decile in its similar schools rank. In contrast, Canyon remained in the first decile in its base score, but did manage to increase from the third to the sixth decile in its similar schools rank.

### *Eastwood District*

Eastwood is the only rural district in our California sample, but like the other California districts, its students represent a wide range of socio-economic and ethnic backgrounds. At the time of our interviews, the superintendent was the only administrator in the district office and Eastwood High was the district’s only high school. Most students at Eastwood are eligible for free- and reduced-price lunch and, at the time of our interviews, the school was not performing well in the state’s accountability system. When we visited the school, Eastwood was undergoing state intervention, which had caught the school by surprise. The relationship between the district and the school was very strained. Many of the teachers noted that there was contention among all of the stakeholders, including the superintendent, the school board, and the teachers. Most teachers and administrators speculated that the superintendent volunteered the school for the external intervention, which brought the principal under fire. Furthermore, the superintendent had fired the assistant principal without any prior warning.

With only one employee, Eastwood District provided little support to its high school. The district had an Assistant Superintendent for Curriculum and Instruction during the 2001–2002 academic year, but when we



interviewed, there was no one providing academic or instructional support to the high school. In a sense, the administrator of the state intervention served some district functions by providing an improvement plan and oversight. However, teachers at the school noted that they felt like this administrator was not providing very much support in implementing the improvement plan.

Despite many challenges, Eastwood's efforts appeared to be starting to improve levels of both internal accountability and alignment with state policies. After a long period of inactivity, Math and English departments had begun to meet regularly to align their curriculum with state standards and had begun to focus on efforts to help students pass the CAHSEE. Consequently, internal accountability was increasing largely as the result of external pressure, as teachers and administrators had begun to build a common instructional culture aimed at addressing the state intervention. The school resisted recommendations from the state intervener that the principal be replaced due to poor student performance. Consequently, while teachers appeared to be aligning, they were doing so in a way that was not completely in line with state mandates. There also appeared to be a strong undercurrent of resistance to external intervention and much of the school's actions vis-à-vis state policies appeared to be motivated by teachers' reluctance to lose their principal. In this sense, it appeared that the district's intervention (by forcing an outside intervention) was only partially successful, as it did not lead to strong support from the school's teachers or administrators for school improvement. Despite the negative feelings toward the intervention, the efforts appeared to have met with at least some success, as the school's 2002–2003 scores on the state API increased one decile on their base score and four deciles on their similar schools rank.

### *Redwood District*

Redwood District is one of two high-school-only districts in our sample. The student population is diverse, with the largest populations being White and Hispanic. Redwood District serves a mix of socioeconomic communities. In general, the district performs well on state accountability measures. Despite this success, the district serves a heterogeneous mix of students with varying levels of academic preparedness, as well as many students who lack basic reading skills. The sample school in this district, Redwood High, also has a heterogeneous mix of students but has been able to draw on some of

the financial resources of some parts of the community. The school had a strong administration with a clear vision and strategies for improving instruction, and the relationship between the district and Redwood High was quite positive.

As a district, Redwood had responded to greater accountability press through efforts to meet the learning needs of English learners via Specially Designed Academic Instruction in English (SDAIE) and other initiatives to differentiate instruction for students. Accountability pressure had also ratcheted up the need for the district to address the test score gap between minority and White students and use standardized test score data to inform schools about individual and departmental strengths and weaknesses. Further, the accountability press increased cooperation between the district and schools. For example, Redwood's English Language Learner department received assistance from a full-time district official charged with overseeing high schools and ensuring that "schools are accountable for better performance of ELL students." The district staff was also undertaking a compliance review audit that monitored the progress of students and reported results to the instructional vice principal at each school. Additionally, the district provided assistance in the form of specialized staff to assist personnel at the high schools. For example, the district's research and evaluation director often worked directly with the data coordinator at Redwood High to strategize and coordinate the dissemination of student data.

While the district had provided assistance in the areas described above, this often came at the urging of school leaders who had taken the lead on initiatives ranging from class size to data management. Often the district provided assistance only after school leaders identified needs and areas for improvement. A district official summed up the role of the district as one of collaboration rather than intrusion. Redwood High also had a dynamic principal, who together with a well-organized school site council, was able to build a strong consensus around instructional goals and strategies. Teachers were working on inter-disciplinary projects such as a senior research project and collaboration regarding how to meet the needs of English learners. The school had also undertaken significant efforts to prepare students for the state tests and was working hard to use student test data to inform instruction. Additionally, the Math department had completed an ambitious project to redesign the Math curriculum and align it with state standards. The effectiveness of these efforts was evident in the school's relatively high 2003 API score and an even higher similar schools rank.

*Urban District*

Urban District has all the problems of many large urban districts: crime, drug use, high student turnover, and high drop out rates. At the time of our interviews, the district had also recently suffered from financial and leadership crises. The population of Urban District is highly diverse. Many schools in the district made large gains in 1999–2002 in their API, which is based mainly on test scores. The high school selected for our sample, Urban High, was among these schools. However, Urban High remained a high school in crisis – an urban school with a large proportion of minority students receiving free or reduced-price lunch, demonstrating low levels of both internal accountability and alignment with state policy. The instructional culture was extremely atomized and teachers rarely worked collaboratively on developing standards-based instruction or any other instructionally focused task. The school appeared to be more focused on day-to-day survival, coping with issues such as security and infrastructure. There seemed to be a reasonably good working relationship between the District and the high school principals, and the principals had a fair amount of autonomy. However, there was some indication from our interviews in the district office that teachers were not buying into efforts to increase test scores. Most of the efforts at school improvement appeared to come from the district rather than from the schools.

Although the district office was responding to accountability, we had a clear sense that they were not enthusiastic supporters of test-based school and student evaluations. The district had also supported a number of other reforms in high schools, such as career academies, converting large high schools into smaller theme schools, and investing heavily in school surveillance and security. The strong accountability press had pushed the district office to begin to introduce new math and reading programs. However, these remained limited and only marginally aligned with accountability requirements.

The district's efforts seemed to have little effect on Urban High's response to external accountability. Faced with difficult challenges on a regular basis, teachers and administrators expressed the idea that testing was unfair for their population of disadvantaged students. For many teachers, this sentiment accompanied indifference toward tests like the SAT9 and CAHSEE. When asked if her department was accountable for passage rates on these tests, the chair of the English department replied, "Not that I know of. You know ... there's controversy about some of those tests." The chair also complained that the accountability press had a demoralizing

effect on many teachers. Another teacher complained that students did not take the SAT9/CAT6 tests seriously, while another expressed doubt that the passage requirements of the CAHSEE would be enforced. Possibly due to this school-wide attitude, in combination with the student population, the school scored very poorly on California's accountability measures in the year that we conducted our interviews.

### *Vine District*

Vine district is a relatively small district located on the outskirts of a large urban area. The district's student population is spread fairly evenly among White, African-American, and Hispanic students. Almost one-third of the students qualified for free- and reduced-price lunch in 2002–2003, and just over one-tenth were English learners. Vine was established as a K-8 district, but with the creation of Vine High in the 1990s, it became a unified district. At the time of our interviews, Vine High was still the district's only comprehensive high school. Although Vine had been scoring in the middle range of California schools on the state's API, our interviews revealed a very low degree of both internal accountability and alignment with state policies. The school's leadership had changed frequently and teachers and departments were extremely atomized. This fragmentation was exacerbated by frequent teacher turnover and the loss of teachers due to budget cuts. While Vine was clearly in need of support from the district, the newness of the school, combined with lack of district experience with high schools, resulted in continued tension and miscommunication between the district and the school.

Due to their historical focus on K-8 schools, officials at Vine District did not have a clear vision of what a high school should look like or do. For example, all of the members of the district's professional development team had elementary school backgrounds and few had any experience with high schools. As one teacher at Vine High noted, "It's not that I don't feel we get support from the district, I just think that the district is still trying to figure out how to help the high school." As a result, the district employed ad hoc and limited efforts to improve and support Vine High. One of the district's most significant changes came as the result of a review after the school's accreditation process, which identified the school's leadership as a problem. In response to this report, the district replaced the principal; the new principal, who had only been at the school for two months when we conducted our interviews, identified weak leadership in the past as one of the

school's key problems. As he explained, "a lot of people had the title ... but they've never taken charge." Other than this major intervention, the district's efforts to encourage a response to state requirements were concentrated primarily on its elementary and middle schools. The district had provided Vine High with preparation materials for the CAHSEE, but otherwise we saw little evidence of a coherent strategy to respond to accountability pressures. For the most part, teachers and school administrators at Vine High were on their own to develop strategies for instructional improvement and to help students meet state standards.

Unable to draw on either strong support from the district or a coherent internal organization, Vine High lacked a clear strategy to respond to state accountability requirements. Additionally, the school's location among the middle band of the state's API rankings weakened the degree of urgency felt by teachers to improve instruction and raise test scores. The school's efforts to align with state accountability policies were limited primarily to motivational events to persuade students to take the state tests seriously. The district's strongest intervention was the recent replacement of the principal. While the new principal had been in place for a very short time, this change may have helped Vine High to develop a stronger sense of its vision and strategies for improvement. It is impossible to attribute test score gains to such an event, but Vine made large gains in the year that we conducted our interviews, increasing two deciles on its API base rank and three deciles on its similar schools rank.

## DISCUSSION

CPRE's strategy to select school pairs within the same district revealed interesting differences in both district improvement strategies and schools' responses to their districts' efforts. Coincidentally, the two districts with two schools each in our sample (Bay and Canyon) were two of the most active districts in the entire sample. Distinct school reactions to district efforts in both of these districts lead us to conclude that schools do not react uniformly to district initiatives. In Redwood District, a high-capacity high school (Redwood High) appeared to take the lead in many of the district's efforts to align with external accountability policies. The remaining three districts in the sample provided very little assistance to their schools in terms of improving internal accountability or aligning with state policies.

Despite these differences, district interventions across our entire sample generally focused on improving schools' alignment with state policies – we

saw very few efforts by districts to help schools increase internal accountability. In fact, districts seem to assume that schools will respond rationally and quickly to their mandates, regardless of schools’ internal organization. However, our empirical results indicating variability in response suggest that schools follow a much more loosely coupled model, as observed by Weick (1976, 1982) and others. We argue that districts must recognize this looseness as well as the need to help strengthen a shared vision in schools, which we have observed to be positively related to school response. In addition to these general observations regarding the districts in our sample, we note the following findings in response to our research questions.

*Are districts responding to state accountability policies and, if so, how do responses vary according to district characteristics?* As Table 2 indicates, our districts vary considerably in their efforts to align their schools with state requirements. While some districts – particularly Canyon (6 out of a total of 6) and Bay (5) – directed strong efforts toward all of their high schools, others like Eastwood (1), Urban (2), and Vine (0) made little or no efforts to encourage their schools to conform to accountability targets. What explains these differences? They could be related to differences in student populations, size, and resources available to schools and districts. Yet, due to our sampling strategy, all of the districts in our sample had high schools with diverse student populations and considerable groups of underperforming and economically disadvantaged students. While we did not see clear patterns along these lines, clear patterns did emerge in terms of other district characteristics. To begin with, districts with a strong high school focus, or at least knowledge of how high schools operate, appear to exert stronger efforts. Two of the three top-scoring districts in Table 2 – Bay (5) and Redwood (4) – only had high schools and presumably understood the difficulties of eliciting responses in these large and complex organizations. In contrast, the two “high-school-hardly” districts in our sample (i.e., each district had only one high school and did not appear to have

**Table 2.** Rating Districts’ Efforts to Align Schools with State Policy.

School	Low-Performing High Schools	Other High Schools	Overall
Bay	3	2	5
Canyon	3	3	6
Eastwood	1	0	1
Redwood	2	2	4
Urban	1	1	2
Vine	0	0	0

a clear or coherent strategy to encourage external response), Eastwood (1) and Vine (0), scored at the bottom of our sample.

These results suggest that in attempting to encourage their schools to respond to state mandates, districts with few high schools and little knowledge of how they operate are at a disadvantage. As Gross and Supovitz (2005) point out, in comparison to elementary and middle schools, high schools tend to be larger and more complex organizations in which teachers have a deep sense of professional autonomy and specialized content knowledge. Moreover, Siskin (2003) argues that by demanding that high schools hold all students to minimum academic standards in order to graduate, recent accountability policies ask high schools to work in contradiction to their historical mission to “serve democratic purposes and accommodate diverse student populations by creating a wide range of programs, and a differentiated curriculum” (p. 177). Our research suggests that districts with little knowledge of these complexities face an uphill battle to encourage their few high schools to change and respond to new requirements.

Paradoxically, the highest-scoring district in our sample, Canyon (6), is a large unified school district with a large assortment of elementary, middle, and high schools. Yet, the district also had the clearest and most comprehensive vision of what its high schools should be like. One of the reasons Canyon scored so highly was that unlike Bay, it had no high schools scoring in the middle range or above on California’s API. Consequently, the district felt strong pressure to improve *all* of its schools, resulting in an all-encompassing strategy, illustrated by strong efforts to increase access to high-quality curriculum for all of its students and provision of extra tutoring and test preparation to prepare students for the CAHSEE. In contrast, Bay District scored slightly lower than Canyon because it differentiated its high school strategy, focusing a set of interventions on schools most at risk of receiving state sanctions.

Of course, Eastwood and Urban also felt strong pressure to increase test scores in their districts. Yet, these districts did not channel this pressure into comprehensive efforts to elicit responses from their schools. In Eastwood, this lack of effort likely resulted from a limited knowledge and vision for its high school, combined with a strained relationship with the high school’s leaders. Urban District, faced with financial and political pressures and other demands associated with being in a crisis mode, simply may have been incapable of developing a coherent strategy to support and encourage its schools to respond to state requirements. Additionally, the district itself did not appear to have a strong sense of internal accountability. While we did

not attempt to measure the degree of internal accountability in districts, the contrast between Urban and Canyon – both large, comprehensive urban districts with many low-performing and economically disadvantaged students – is striking. Canyon’s clear sense of vision and comprehensive strategies placed it at the top of our district sample, while Urban is next to last of the six districts.

Considering the findings related to Research Question 1 through the lens of organizational theory, it appears that the most successful districts were those that recognized the complexity of the high school organization and had a coherent vision of improvement. However, most of the district actions across the sample applied a rational bureaucratic model in their responses to accountability (as suggested by Weick, 1982). The districts provided support for certain activities or professional development with the expectation that these interventions would result in the changes and responses they sought. This may have influenced instruction simply because the district allocated resources to these activities (see Gamoran & Dreeben, 1986). However, the districts in our sample generally did not engage in activities aimed at strengthening internal accountability or a common vision. Thus, it appeared that districts did not fully recognize the loose coupling between the school and the district and between school administrators and teachers, or at least did not respond in ways that appeared to recognize this fundamental aspect of the organization.

*How do district responses vary according to school characteristics?* Of the districts in our sample that differentiated their approaches based on the characteristics of their schools, the strongest determinant of differentiation appeared to be the performance of schools relative to the expectations of the state. That is, these districts focused the greatest efforts on the schools most likely to receive state sanctions due to low student performance. For example, while Bay District had a set of common interventions for all high schools such as biannual testing of all students, the district had an additional set of strategies only for schools identified by the state as low performing. But we also note that some districts, like Canyon, could not employ a differentiated strategy simply because all of their high schools were in the low-performing group. Consequently, Canyon District had a strong set of interventions for all of its high schools. On the other end of the student performance spectrum, Redwood District, which appeared to consider itself a high-performing district, also directed relatively strong and undifferentiated efforts toward all of its schools. Yet, the lower pressure on this district resulted in lower overall effort relative to districts like Bay and Canyon.



Ironically, two of the districts facing the strongest pressure to raise student performance had the lowest degree of effort (Urban and Eastwood). We attribute this low effort to the overall environment of crisis in the districts and their low-performing schools. These districts seemed to have much more to concern themselves with and simply could not organize themselves adequately to mobilize the same degree of effort as the other districts. As a contrast, we believe that while Canyon had the potential to be a district in crisis (certainly Canyon High School was facing crisis conditions), a strong district leadership team and vision placed it instead at the top of our sample. The lowest-scoring district in our sample, Vine High, featured a combination of relatively low pressure – due to a high school that was not scoring in the bottom-two deciles of California schools at the time of our interviews – and little knowledge or experience working with high schools. These findings once again underscore the loose coupling at every level of the education system – from states to districts, and from districts to schools. While accountability policies are aimed at ensuring that all students succeed, many schools that are below average do not demonstrate a pressing need for improvement.

*What is the relationship between district efforts and schools' responses?* Schools' responses to district efforts ranged from indifference to active cooperation with district initiatives. We did not see evidence of schools pushing back or actively resisting district efforts. One exception was Eastwood High, whose teachers felt that the district was undermining the school by firing well-liked administrators at inopportune times. We attribute Eastwood's strained relationship with its district to a combination of non-existent district support and the school's own lack of internal accountability. Fig. 3 demonstrates the relationship between districts' efforts to align their schools with state mandates and actual school responses. There appears to be some relationship between the strength of districts' efforts and responses of the schools. For districts placed above the *x*-axis in terms of their efforts (Canyon, Bay, and Redwood), corresponding schools rank above the others in terms of external response. Yet, there are also exceptions to this relationship. In many cases, there is a sizeable difference between district and school efforts. The difference between the district and the school in terms of external efforts, or the response gap, may be considered as the failure of district efforts to result in school responses. By our illustrative coding scheme, the largest such gap occurs between Canyon District and Canyon High (–4). There are also substantial gaps between Bay District and both of its high schools (–3) and between Urban District and Urban High (–2). There are also a few cases in which school responses actually

exceed district efforts, as in Vine High (+2) and Redwood and Eastwood Highs (+1). What explains these response gaps?

We conclude that two potential determinants of response gaps are a school's internal accountability and the pressure it feels from the state. To begin with, in a district like Canyon in which all schools feel strong pressure from the state, the school's internal accountability influences the school's ability to respond to this pressure. Despite the district's strong efforts, Canyon High's low degree of internal accountability (2 on the internal accountability scale) impeded the school's effort to respond strongly to pressure. In contrast, San Antonio High, which had a very similar student population to that of Canyon, had a much greater degree of internal accountability, with a score of 8. The school's strong shared sense of vision and collaboration enabled it to mobilize a powerful response to state requirements. In fact, the relationship between internal accountability and external response is fairly strong across all eight of our schools, as indicated by the upward trend of the data points in [Fig. 3](#). Seemingly, even with strong district efforts, schools without strong internal accountability have difficulty responding to state pressure.

Conversely, without strong pressure from the state, even schools with stronger internal accountability may not have the incentive to respond to district mandates. For example, of the two schools in Bay District, Pinewood (6) scored higher in internal accountability than Arnold (3). Yet, we gave each school a 2 in terms of external response. While Pinewood's stronger internal accountability could have resulted in a stronger response to state requirements, the school was scoring high enough on the state API not to feel a sense of urgency about "getting the scores up." And despite strong pressure from the state, schools like Urban and Eastwood simply did not have strong enough internal accountability to convert that pressure into action. A muted response from these schools' districts reinforced this lack of response. Furthermore, many of the districts in our sample did not appear to address the very nature of the loosely coupled systems they were dealing with. While they may have provided direct services to schools, they did little to address the underlying ability of school actors to respond to the accountability system.

Overall, we conclude that while districts can support schools' efforts to align with state policies in many ways, a lack of strong internal accountability at the school level may result in a large response gap between the district and the school. This gap occurs in part due to district difficulties in effecting change when facing the loosely coupled organizational structure of schools. While district officials may recognize the looseness of ties among

school actors and events, most of their efforts assume a tighter organizational framework of action and response. Based on our observations, very few district efforts are directed toward strengthening common shared visions within schools. Yet, this common vision, or internal accountability, appears to be strongly related to schools' ability to respond. Distinct responses from schools in the same district, such as San Antonio and Canyon, illustrate how district interventions can meet with different responses in schools with varying levels of internal accountability. This suggests that in addition to encouraging schools to engage in activities like test preparation, extra tutoring for struggling students, and use of student test score data to inform and focus instructional efforts, districts must also recognize the importance of school organization and seek ways to support stronger internal accountability in their schools. More generally, our observations underscore the importance of understanding organizational behavior when designing policies and approaches to school reform.

What does our work tell us about the behavior of schools and districts as organizations? To begin with, our results (albeit from a limited sample) support the work of earlier researchers who have argued that schools are loosely coupled in terms of both internal actors and external relationships. In particular, the observations of Weick (1982) regarding the weak, unpredictable, and intermittent nature of ties among people ring true when applied to our schools and districts. The looseness of ties among people, as well as between mandates and action, leads to large response gaps and frustration on the part of those issuing mandates. These gaps are most evident in examining differences between seemingly clear directives from states and districts to improve instruction of low-achieving students and the lack of direct school efforts to do so. As Weick (1982) observed, when faced with such inconsistencies, stakeholders apply the "rational bureaucratic model" to their understanding of the situation and conclude that someone must be blamed or fired. In fact, leaders at two of our sample schools were fired when district officials came to feel that their efforts to improve instruction were insufficient.

Of course, inconsistencies between mandates and actions do not mean that districts *cannot* influence instruction. As Gamoran and Dreeben (1986) have argued, districts can allocate resources to influence classroom instruction. Weick (1982) has argued for greater efforts to organize instruction around a common vision and key values. In either case, the influence of external and internal actors is likely to be both indirect and unpredictable. While our analysis attempts to identify factors that predict greater and more aligned response, such as internal accountability and immediacy of

sanctions, there are undoubtedly many other unidentified variables that influence school and teacher actions.

What if our conception of schools as loosely coupled systems is mistaken and U.S. schools actually follow a rational bureaucratic model? We would be more likely to see a clear and immediate reaction at all levels of the system. In addition, the district efforts we observed, which were largely aimed at increasing external response, would meet with more uniform success. As a consequence, well-designed district mandates alone may be sufficient to elicit parallel responses. But as we have observed, schools need much more – including the development of internal accountability – to respond to state and district mandates as expected.

While there is no clear strategy for developing a shared vision in complex organizations like high schools, our interviews in California schools and districts suggest that one key may be to focus on developing strong school leadership teams. A few of the districts in our sample resorted to the drastic solution of sacking administrators with little warning, but our research indicates that strong leadership extends beyond the principal's office to support staff, department chairs, and teachers themselves. Our results suggest the importance of developing and supporting "distributed leadership" (Spillane, Halverson, & Diamond, 1999), in which school leaders emerge from across the school and take on various yet related roles.

From Weick's (1982) perspective, in loosely coupled systems like schools, leadership is not concentrated in a single person or office. Instead, due to a variety of local initiatives, the amount of leadership tends to be greater but more unfocused than in a tightly coupled system. But if we conceive of schools as loosely coupled systems, we must also recognize the importance of focusing this leadership around "strong collective views of what they stand for, and well-developed organizational processes that bring those beliefs into action" (Elmore, 2003, p. 196). Of the high schools in our sample, San Antonio High best exemplified this leadership model. At the same time, we also found that the presence of a strong principal with a clear vision and plan for raising student achievement can also augment a school's internal accountability. The importance of this type of leadership was particularly evident in Redwood High, a school that, led by a strong principal, often influenced the district's actions toward meeting the state's requirements.

Internal accountability may be a necessary condition for a school to respond to state requirements, but it also must be accompanied by strong motivation to improve student performance. This was particularly clear in Pinewood High, which had scored in the middle range of California's API

and considered its status in this range as a reason to pay little heed to state requirements like the impending CAHSEE exam. While one could argue that such schools have the “right” to ignore state mandates, Pinewood performed poorly relative to schools with similar student populations. To encourage improvement efforts in such schools, districts must develop ways to redefine or focus state requirements in a way that provides real motivation to raise student performance.

Finally, districts seeking to encourage strong responses in schools can focus on building schools’ resources, knowledge, and skills to develop and enact strategies to improve student performance. As [Weinbaum \(2005\)](#) explains, districts’ efforts to do so can take many forms, including professional development, encouraging effective use of student test score data, and allocating resources effectively. While such efforts cannot guarantee that schools will use them, particularly in schools where internal accountability is weak, they will provide schools with the tools they need once they develop a coherent vision and strategy to improve.

## NOTES

1. Some scholars have looked at outlier districts – those with unexpectedly high levels of student performance or reputation for bringing about impressive change – to examine what actions these districts have taken (see e.g., [McLaughlin & Talbert, 2003](#)). However, these districts do not represent the majority of districts nationwide ([Weinbaum, 2005](#)).

2. These designations are based on earlier work by [Goertz and Duffy \(2001\)](#) and [Carnoy and Loeb \(2002\)](#), which characterized states according to the strength of sanctions for both schools and students. Our frame of reference for accountability systems in this chapter is the academic year 2002–2003. For a more detailed description of the sampling strategy, see [Gross and Supovitz \(2005\)](#).

3. Obvious problems arise with the validity and reliability in using such a rating system to quantify complex concepts like internal accountability and policy response. Even more problematic is the assumption, implicit in our later analysis, that differences between ratings represent real and quantifiable gaps among schools and between schools and districts. Our attempt to quantify our findings stems from our desire to continue and make more systematic a useful heuristic originating in the work of [DeBray et al. \(2003\)](#). In particular, we believe that the graphical representations generated from the numeric ratings system provide a powerful method to compare schools and identify potential relationships. However, in recognition of the limitations of our quantification scheme, we complement this analysis with a narrative treatment of our findings. We also encourage other researchers to critique and refine our methods to compare and understand the complex actions of schools and districts.

4. District and school names are pseudonyms.

## ACKNOWLEDGEMENTS

The research for this chapter was conducted as part of a larger effort by the CPRE and funded by the Institute of Education Sciences, United States Department of Education. Opinions in this chapter are our own and do not necessarily reflect the views of the Institute of Education Sciences, the United States Department of Education, or CPRE. We would like to thank Martin Carnoy and anonymous reviewers for their helpful comments on various drafts of this chapter. Any errors are of course our own.

## REFERENCES

- Abelmann, C., & Elmore, R. (1999). *When accountability knocks, will anyone answer?* CPRE Research Report Series RR-42. Consortium for Policy Research in Education, Philadelphia, PA.
- Adamson, F., Carnoy, M., Addy, N., Ricalde, B., & Rhodes, L. (2007). *Data use, English language learners, and literacy programs in California high schools: A principal survey*. Unpublished manuscript, Stanford University.
- Carnoy, M., Elmore, R., & Siskin, L. S. (Eds). (2003). *The new accountability: High schools and high-stakes testing*. New York: RoutledgeFarmer.
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305–331.
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools*. Washington, DC: The Brookings Institution.
- DeBray, E., Parson, G., & Avila, S. (2003). Internal alignment and external pressure: High school responses in four state contexts. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *The new accountability: High schools and high-stakes testing* (pp. 55–85). New York: RoutledgeFarmer.
- Elmore, R. (2003). Accountability and capacity. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *The new accountability: High schools and high-stakes testing* (pp. 195–209). New York: RoutledgeFarmer.
- Firestone, W. A. (1985). The study of loose coupling: Problems, progress, and prospects. In: A. Kerckhoff (Ed.), *Research in sociology of education and socialization* (Vol. 5, pp. 3–30). Greenwich, CT: JAI Press.
- Fuhrman, S., & Elmore, R. (1990). Understanding local control in the wake of state educational reform. *Educational Evaluation and Policy Analysis*, 12(1), 82–96.
- Gamoran, A., & Dreeben, R. (1986). Coupling and control in educational organizations: An explication and illustrative comparative test. *Administrative Science Quarterly*, 31, 612–632.
- Goertz, M. E., & Duffy, M. C. (2001). *Assessment and accountability systems in the 50 states: 1999–2000*. Philadelphia, PA: Consortium for Policy Research in Education.
- Goertz, M. E., Massell, D., & Chun, T. (1998). *District response to state accountability systems*. Paper presented at the annual meeting of the Association for Public Policy Analysis and Management, October, New York.

- Gross, B., & Supovitz, J. A. (2005). Introduction. In: B. Gross & M. E. Goertz (Eds), *Holding high hopes: How high schools respond to state accountability policies* (CPRE Research Report No. RR-056, pp. 1–16). Philadelphia, PA: Consortium for Policy Research in Education.
- Luschei, T. F., Goldwasser, M., Carnoy, M., Christensen, G. S., & Weinbaum, E. H. (2007). *Internal accountability and external alignment: Evidence from forty-one high schools in five states*. Unpublished manuscript, Stanford University.
- March, J. G., & Olsen, J. P. (Eds). (1976). *Ambiguity and Choice in Organizations*. Bergen, Norway: Universitetsforlaget.
- Marsh, J. S., Kerr, K. A., Ikemoto, G. S., Darilek, H., Suttorp, M., Zimmer, R. W., & Barney, H. (2005). *The role of districts in fostering instructional improvement. Lessons from three urban districts partnered with the Institute for Learning*. Santa Monica, CA: RAND Corporation.
- Massell, D., Goertz, M., Christensen, G., & Goldwasser, M. (2005). The press from above, the pull from below: High school responses to accountability. In: B. Gross & M. E. Goertz (Eds), *Holding high hopes: How high schools respond to state accountability policies* (CPRE Research Report No. RR-056, pp. 17–35). Philadelphia, PA: Consortium for Policy Research in Education.
- McDermott, K. A. (2006). Incentives, capacity, and implementation: Evidence from Massachusetts education reform. *Journal of Public Administration Research and Theory*, 16(1), 45–65.
- McLaughlin, M., & Talbert, J. (2003). *Reforming districts: How districts support school reform*. Seattle, WA: Center for the Study of Teaching and Policy.
- Meyer, J. W. (1980). Levels of the educational system and schooling effects. In: C. E. Bidwell & D. M. Windham (Eds), *The analysis of educational productivity* (Vol. 2, pp. 15–63). Cambridge, MA: Ballinger.
- Meyer, J. W. (1986). Types of explanation in the sociology of education. In: J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 341–359). Westport, CT: Greenwood Press, Inc.
- Meyer, J. W., & Rowan, B. (1978). The structure of educational organizations. In: M. W. Meyer (Ed.), *Environments and organizations* (pp. 78–109). San Francisco, CA: Jossey-Bass.
- Meyer, J. W., Scott, W. R., & Deal, T. E. (1983). Research on school and district organization. In: J. V. Baldrige & T. Deal (Eds), *The dynamics of organizational change in education* (pp. 409–425). Berkeley, CA: McCutchan.
- O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72(3), 293–329.
- Orton, J. D., & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review*, 15(2), 203–223.
- Siskin, L. S. (2003). When an irresistible force meets an immovable object. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *The new accountability: High schools and high-stakes testing* (pp. 175–194). New York: RoutledgeFalmer.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. *Politics of Education Association Yearbook 1990*, pp. 233–267.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (1999). Investigating school leadership practice: A distributed perspective. *Educational Researcher*, 30(3), 23–28.
- Weick, K. E. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1–19.

- Weick, K. E. (1982). Administering education in loosely coupled schools. *Phi Delta Kappan*, 63(10), 673–676.
- Weinbaum, E. H. (2005). Stuck in the middle with you: District response to state accountability. In: B. Gross & M. E. Goertz (Eds), *Holding high hopes: How high schools respond to state accountability policies*. CPRE Research Report No. RR-056, pp. 95–121. Philadelphia, PA: Consortium for Policy Research in Education.
- Wilson, B. L., & Corbett, H. D. (1983). Organization and change: The effects of school linkages on the quantity of implementation. *Educational Administration Quarterly*, 19(4), 85–104.
- Woody, E. L., Bae, S., Park, S., & Russell, J. (2006). *Snapshots of reform: District efforts to raise achievement across diverse communities in California*. Berkeley, CA: Policy Analysis for California Education.



# CHAPTER 4

## TIGHTENING THE SHIP OR SLOWLY SINKING? RESHAPING TEACHERS' WORK CONDITIONS

Kristin Gordon

### ABSTRACT

*The most recent development in the accountability movement occurred in January 2002 when the No Child Left Behind Act (NCLB) was signed into law. Surprisingly little work has illuminated how teachers experience standards-based accountability policy. Using survey data and interviews, this chapter explores the impact of NCLB requirements, namely adequate yearly progress and needs improvement status, on teacher perceptions of working conditions, especially the use of time and empowerment. I show how the policy has led to restructuring of classroom time and increases in collaboration and yet, simultaneously, a decrease in teachers' perceptions of empowerment.*

School accountability and standardization have dominated the United States reform discourse for the past quarter-century. By the mid-1980s, most states had developed their own accountability plans in response to this growing trend. While the Clinton administration attempted to formalize these state plans with limited success, a dramatic shift in federal involvement

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
**Research in Sociology of Education, Volume 16, 103–131**  
Copyright © 2008 by Emerald Group Publishing Limited  
All rights of reproduction in any form reserved  
ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16004-9

in educational accountability occurred in early 2002 when President George W. Bush signed the No Child Left Behind Act (NCLB) of 2001 (NCLB, 2002). Arguably, not since the Elementary and Secondary Education Act of 1965 has the United States government taken such a bold step towards influencing the educational process.

The federal NCLB follows and punctuates earlier state-led accountability efforts, going back to the mid-1980s. This law centralized accountability by setting national objectives, mandating how achievement should be tracked locally, and sanctioning schools that fail to show gains. In an effort to increase achievement across all students, a nationwide push to reach 100% proficiency in reading and math by 2014 became the new goal of the federal government. States were charged with the task of developing academic standards and annual tests to chart their progress. NCLB raised the stakes of accountability demands through the use of federally mandated sanctions. Under this law, schools and school districts that are unsuccessful at meeting academic expectations (Adequate Yearly Progress or AYP) for two consecutive years are identified as Needs Improvement schools and become subject to a series of sanctions. Sanctions range from school choice options and supplemental services for students to state guided contract monitoring, replacement of school staff, and eventual school reconstitution (NCLB, 2002). Through these policy mechanisms, NCLB attempts to penetrate into the internal operations of schools, a territory historically dominated by the state and local education agencies. Unlike previous policy, which focused primarily on compositional issues such as school segregation and funding, NCLB goes straight to the bottom line of education by requiring measurable changes in student achievement.

Schools are not just places to educate children. They also represent workplaces for teachers and administrators. To understand the impact of top-down accountability policies, we must study the effect of these changes on teachers' working conditions. NCLB holds teachers accountable for the academic performance of students by directing many of the initiatives, requirements, and even some of the sanctions at them. In this way, examining teachers and their working conditions is especially important given their centrality to the accountability movement. While considerable research has examined the effect of accountability policies on student achievement and equality, few investigations have focused on teachers' working conditions. This chapter seeks to fill this gap by examining how NCLB requirements, namely needs improvement status and AYP pressures, affect teachers' perceptions of their working conditions.<sup>1</sup> This study employs two complementary sources of data. First, with analyses of survey data, the chapter investigates the statistical relationship between

needs improvement status, AYP, and teacher reports of their perceptions of time use and empowerment in their school. Second, the chapter draws on transcripts from semi-structured interviews conducted with teachers in three schools at different needs improvement levels to expand on the quantitative findings. Careful inspection of these vivid descriptions of teachers' work experiences shows that NCLB requirements directly penetrate into American classrooms and adjust the work of teachers in ways they see as both positive and negative.

## **SCHOOL STRUCTURE AND THE AIMS OF ACCOUNTABILITY**

Top-down accountability policies aim to raise academic achievement to a minimum standard and reduce the achievement gap between advantaged and disadvantaged students by holding schools and school districts accountable. States then use differing measures of student performance to evaluate the effectiveness of schools and local educators. These systems rely on students' scores to signal the alleged effectiveness of schools (NCLB, 2001; Hoffer, 2000). If test scores do not demonstrate sufficient academic achievement for all students, then the solution is to "tighten the ship" through sanctions of rising level of magnitude (Ingersoll, 2003, p. 35).<sup>2</sup>

This approach seems logical given the highly decentralized and complex structure of schooling in the United States. The structure of schooling is based on the need to efficiently educate large numbers of children. To complete this task, rationalized bureaucracy would seem the best organizational form from an administrative viewpoint (Weber, 1946; Bidwell, 1965; Ingersoll, 2001). With this organizational structure, academic tasks and assessments must be standardized and the work of teachers needs to be regulated. However, some research on the structure of schools in the U.S. context suggests that this vision of school organization is at odds with the perspective of many teachers. According to several of the classic works on teaching, the practice of educating children does not always lend itself to highly regulated processes (Lortie, 1975; Waller, 1932; Bidwell, 1965). In this view, teaching requires fluidity. Inside the classroom, teachers are thought to be most effective when they operate with relative autonomy, adapting to the differing needs and demands of individual classrooms and students. At the same time, research suggests that the structure of schools in the United States tends to isolate teachers in classrooms. Lortie (1975, p. 23)

states, “the subsequent work relationships of teachers have been marked by more separation than by interdependence; most teachers still spend most of their time working alone with a group of students in a bounded area.” Thus, the American system seems to be composed of two somewhat contradictory structural needs, that of bureaucratic regulation by administrators and the pursuit of autonomy by teachers.

To resolve this contradiction, [Bidwell \(1965\)](#) suggests that schools have developed a certain amount of structural looseness to allow teachers and schools to exercise autonomy while the district and state systems continue to be hierarchically organized. Considerable research has identified schools as loosely coupled systems (for example, [Elmore, 2000](#); [Meyer & Rowan, 1978](#); [Dreeben, 1973](#)). This contradiction in educational organization has created continual debate over the “correct” description of U.S. school structure.

The “looseness” means that one’s perspective of the organization of U.S. schools is determined by one’s own position in the system. From an administrative standpoint, where schools are compared to a rationalized bureaucracy, schools appear disorganized. [Ingersoll \(2003\)](#) calls this perspective the school disorganization perspective. From this viewpoint, increasing control over school processes can rectify academic and pedagogical problems that originate in schools. In other words, this perspective suggests that “tightening the ship” will address the problems of education ([Ingersoll, 2005, 2003](#); [Rowan, 1990](#)). For example, current government initiatives to increase teacher and school accountability through the use of standardized curriculum, high-stakes testing, and sanctions all respond to the perception that teachers are overly autonomous in their classrooms.

While the administrative viewpoint sees schools in the United States as structurally disorganized, teachers hold a different view of school organization. From teachers’ vantage point, called the disempowerment perspective, schools seem to be ideal examples of hierarchical, bureaucratically controlled organizations. Teachers perceive themselves to be highly regulated and subject to top-down authority structures, and thus disempowered ([Ingersoll, 2005, 2003](#); [Rowan, 1990](#)). For example, while serving their students, teachers also face administrative responsibilities, now intensified by growing pressure from the sanctions associated with accountability ([Sunderman, Kim, & Orfield, 2005](#); [Costigan & Crocco, 2004](#)). In contrast to the disorganization perspective, the disempowerment perspective suggests further decentralizing control over school processes will empower teachers to customize education to the needs of students

and thereby raise the academic achievement of all students. Thus, while the aims of policy makers, administrators, and teachers are similar, their perspectives and solutions vary according to their organizational position.

The highly decentralized structure of schooling in the United States resulting in this seeming contradiction is not shared in many national contexts (e.g. Baker & LeTendre, 2005; Fuller & Rubinson, 1992). In fact, the decentralized history of U.S. schools creates the structural space for the development of differing perspectives by administrators, policy makers, and teachers. These differing vantage points and their structural roots are important when examining how teachers experience shifting working conditions under NCLB.

## **ACCOUNTABILITY POLICY AND TEACHERS' WORKING CONDITIONS**

The unique organizational structure and workplace culture of schools make five distinct working conditions crucial to the work of teachers: time, facilities and resources, empowerment, leadership, and professional development (Hirsch, Emerick, Church, & Fuller, 2006; Ingersoll, 2003). This chapter focuses empirically on two working conditions: time use and professional empowerment. These conditions capture distinct elements of teachers' work experiences. Time availability and how it is allocated represent core features of the daily mechanics of teachers' work. In this study, three components of teachers' time are assessed: time on student learning, time on administrative duties, and time spent working with colleagues. Empowerment captures teachers' views of the character of their involvement and participation in decision-making and curricular development. Empowerment is assessed through teachers' reported involvement in school-wide action and decision-making; the level of trust, respect, and expertise in curricular development and implementation; and reported levels of collaboration.

### *Time*

Research demonstrates that teachers in the United States spend more time in front of the classroom teaching or controlling the classroom environment than teachers in many other nations (Schaub & Baker, 1991; Hirsch et al., 2006; OECD, 2003). Time is a scarce resource for United States teachers. They need time to teach the curriculum, plan lessons, and attend to their

own professional development. In addition to these teaching tasks, a considerable amount of time is consumed by duties such as bus and lunch patrol, paperwork, and committee participation. Time availability is important to teachers as they evaluate their working conditions. Many teachers perceive that more time for instructional tasks, collaborative efforts, and planning would improve student learning. The use of time is particularly salient in the current policy context. Teachers report that as a result of state mandated testing, more time is being spent on tested subjects and students close to proficiency than on non-tested subjects and students already proficient or severely behind (Sunderman et al., 2005). As NCLB requirements and sanctions tighten the relationship between the actual work of teachers and accountability expectations, it is probable that teachers will report a restructuring of their time. This direct impact on the work of teachers is measured by their perceptions of the use and availability of time across differing kinds of schools, including whether the school is meeting AYP growth targets or has been pushed into needs improvement status.

### *Empowerment*

External accountability policies may influence teachers' sense of empowerment within their classrooms and in schools. Prior research on teacher control over classroom and school decisions has shown that teachers tend to be relatively autonomous in their classrooms; however, they feel very little control over the larger institutional decisions affecting their work (Ingersoll, 2003; Hirsch et al., 2006). Ingersoll's (2003) work on teacher control demonstrates that there is reduced turnover and more teacher engagement in schools where teachers feel empowered to make key workplace decisions. Teachers want to feel respected as professionals, appreciated, valued, and important for the operation of the school system. Empowerment is not just about making decisions, but also about feeling trusted and having trust in the school leadership (Bryk & Schneider, 2002). Teachers' perceptions of their empowerment appear to be shaped by the extent to which they feel important to the operation of the school and the education of their students.

The National Center for Education Statistics (1997) found a connection between four aspects of teacher professionalism and career commitment. The four aspects of professionalism were teacher classroom autonomy, faculty policymaking influence, assistance for new teachers, and maximum end-of-career salaries. These findings further support the notion that

teachers derive satisfaction from feeling empowered in their classroom and having an influence on larger school issues through school-wide collaboration and involvement. NCLB policy may reshape several of these aspects that lead to higher teacher commitment. With the increased federal and state control over school processes required by NCLB, the nature of teacher empowerment is a particularly relevant aspect of working conditions to investigate. In the U.S. decentralized context, NCLB accountability requirements may reduce the perception of teacher empowerment in the classroom and increase the amount of external control over school decision-making.

## **SURVEY DATA ANALYSIS: TEACHERS' TIME AND PROFESSIONAL EMPOWERMENT**

In January 2005, the Georgia Quality Learning and Teaching Environments Initiative (QLTE) conducted a study of the working conditions in ten public school districts throughout Georgia.<sup>3</sup> The QLTE survey measures educators' level of agreement with statements about specific components of the five domains of working conditions – time, facilities and resources, leadership, empowerment, and professional learning. For each domain, a series of statements assess various components of the working condition and record the educator's level of agreement (1 = "strongly disagree" to 5 = "strongly agree"). Following extensive analyses on each of the survey items, a set of statements from each domain was selected and combined to form a single index gauging the adequacy of that working condition for the respondent.<sup>4</sup> These measures of time and empowerment serve as dependent variables in the OLS regression analysis presented in [Table 2](#).

There are two different ways of examining the relationship between NCLB requirements and teachers' assessments of school working conditions. First, one can think of NCLB sanctions as having a cumulative impact on teachers. If this is the case, then one would need to measure sanctions using the number of years in needs improvement status.<sup>5</sup> The assumption underlying this approach is that NCLB intensifies over time and the prolonged exposure to sanctions may produce distinct interpretations by teachers of their working conditions. However, it may be the case that the immediate impact of the failure to meet AYP standards in the previous year is more salient for teachers' experiences. From this perspective, the long-term effect of NCLB sanctions would have a lesser impact than the

result of the last years attempt to meet AYP standards. To capture the distinct effects of NCLB requirements, both measures of the policy are analyzed.<sup>6</sup>

To control for other school and individual level factors that may influence teachers' evaluation of time and empowerment. This analysis includes measures of school characteristics provided by the Georgia Department of Education as well as teacher demographics collected on the QLTE. The control variables include teacher years of experience, race, gender, grade level, school Title 1 status, percent of minority students, students with disabilities, students with limited English proficiency, and faculty size. The characteristics of schools are included as attributes of individual teachers. A total of ten Georgia school districts participated in the QLTE; however, a technical error prevented the identification of schools in three of the participating counties. Thus, this analysis includes all teacher respondents in the seven counties for which complete data is available ( $N = 3,214$ ) (Table 1).<sup>7</sup>

Prior to presenting the regression analysis exploring the relationships among time, empowerment, and the policy variables, a few comments on the impact of the control variables are appropriate. As seen in Table 2, teachers' assessments of time are shaped by all of the control variables with the exception of grade level. Years of teaching experience and gender are positively related to perceptions of availability and use of time. The percent of students with disabilities, minority enrollment and faculty size lowered perceptions of the availability and use of time. Surprisingly, race, Title 1 status, and limited English proficiency are positively related to teachers' perceptions of time.

For empowerment, teachers' evaluations of empowerment are shaped by all of the control variables with the exception of gender and percentage of limited English proficiency students. Years of teaching experience are positively related to perceptions of empowerment. Grade level, Title 1 status, percent students with disabilities, percent minority enrollment, and faculty size lowered perceptions of teacher empowerment. Interestingly, African American teachers tend to report more empowerment than teachers of other racial and ethnic backgrounds.<sup>8</sup>

### *Time – Teacher Survey Results*

The regression models indicate that the number of years in needs improvement status is associated with a positive assessment of the adequacy of teacher time. In other words, the longer a school has been in needs



**Table 1.** Variables and Descriptive Statistics.

Variable Name	Operationalization	Mean	Standard Deviation	Min	Max
<i>Dependent Variables</i>					
Time	<p>Single index of time includes the following items:                      Teachers have class sizes which afford them time to meet the educational needs of all students.                      Teachers have student loads which afford them time to meet the educational needs of all students.                      Teachers have time during the school day to collaborate productively with their colleagues.                      The school leadership works to reduce routine administrative duties and paperwork so teachers can focus on educating students.                      1 = Strongly Disagree, 2 = Somewhat Disagree, 3 = Neither Disagree or Agree, 4 = Somewhat Agree, 5 = Strongly Agree                      Scale reliability coefficient: <math>\alpha = 0.76</math></p>	11.20	4.27	4	20
Empowerment	<p>Single index of empowerment includes the following items:                      School administrators involve teachers in developing and implementing the school improvement plan.                      Teachers are recognized as educational experts and trusted to make sound professional decisions about instruction and student progress.                      Teachers and staff at my school feel empowered to try new instructional approaches.                      There is an atmosphere of trust and mutual respect at my school.                      I feel comfortable raising issues and concerns which are important to me.                      Teachers and staff work together to improve teaching and learning.                      Teachers help establish and implement policies for student learning.</p>	34.50	8.39	9	45

*Table 1. (Continued)*

Variable Name	Operationalization	Mean	Standard Deviation	Min	Max
	Teachers play a role in the development of the school professional learning plan. A sustained effort is made in my school to empower teachers and parents and other members of the school community. 1 = Strongly Disagree, 2 = Somewhat Disagree, 3 = Neither Disagree or Agree, 4 = Somewhat Agree, 5 = Strongly Agree Scale reliability coefficient: alpha = 0.93				
<i>Independent Variables</i>					
Needs improvement	Number of years in needs improvement	0.53	1.34	0	6
AYP status	Dummy: 1 = failed to meet AYP, 0 = met AYP	0.29	0.46	0	1
Experience	Total years of teaching experience 1 = 1 to 2 years, 2 = 3 to 5 years, 3 = 6 to 12 years, 4 = 13 to 18 years, 5 = 19 to 25 years, 6 = more than 25 years	3.64	1.52	1	6
Race	Dummy: 1 = Black, 0 = non-Black	0.21	0.41	0	1
Gender	Dummy: 1 = male, 0 = female	0.15	0.36	0	1
Grade level	1 = elementary school, 2 = middle school, 3 = high school	1.64	0.82	1	3
Title 1	Dummy: 1 = yes, 0 = no	0.61	0.49	0	1
Disabilities	Percentage of students with disabilities	12.24	4.27	0	22
Limited English proficiency	Percentage of students with limited English proficiency	1.50	2.58	0	18
Minority enrollment	Percentage of minority enrollment	68.19	26.31	0	100
Faculty size	Total number of full time teachers	43.96	20.53	11	146

**Table 2.** OLS Estimates for a Model Predicting Teachers' Time and Empowerment.

	Time	Empowerment
<i>Policy Variables</i>		
Needs improvement	.187** (.074)	.342** (.151)
AYP	-.124 (.225)	-1.879*** (.494)
<i>Control Variables</i>		
Experience	.203*** (.051)	.453*** (.103)
Race (1 = African-American)	1.024*** (.188)	1.930*** (.375)
Gender (1 = Male)	.795*** (.231)	.335 (.497)
Grade level	.030 (.154)	-.768** (.334)
Title 1	1.113*** (.222)	-.811* (.443)
Disabilities	-.043** (.022)	-.129** (.042)
Limited English proficiency	.0987** (.036)	.080 (.061)
Minority enrollment	-.015*** (.004)	-.048*** (.009)
Faculty size	-.0115** (.005)	-.041** (.013)
Constant	10.858 (.531)	40.452 (1.040)
$R^2$	0.041	0.061
$N$	2897	2812

Notes: Robust standard errors reported in parentheses. \* $p < .10$ ; \*\* $p < .05$ ; \*\*\* $p < .001$  (two-tailed tests).

improvement status, the more positively teachers in that school evaluate the adequacy of their time for meeting student needs, completing administrative duties, and collaborating with their colleagues. This finding, while consistent with the objectives of NCLB, may appear surprising to those who would expect the cumulative impact of needs improvement and the accompanying sanctions to have a negative impact on teachers' time.

When NCLB is measured using AYP failure in the preceding year, no significant relationship emerges between AYP status and time. Thus, the impact of accountability policy on time is captured in the long-term effects

of needs improvement, rather than in the year-to-year shift or short-term effects of AYP failure. Overall, there is a general improvement of teachers' perceptions of the adequacy of time as schools move forward with the needs improvement process. This set of findings implies that the duration of needs improvement status and the accompanying sanctions have a positive cumulative effect on one component of teachers' working conditions – time (see Table 2).

### *Perceived Empowerment – Teacher Survey Results*

If the cumulative impact of NCLB requirements on teacher's perceptions of time is positive, one might expect to find that the same pattern exists in teachers' assessments of their empowerment. In fact, this is the case: the number of years in needs improvement is positively associated with teachers' sense of empowerment. Surprisingly, the previous year's AYP failure is also significant. In this case, however, AYP failure is linked to a reduction in teachers' empowerment. Thus, both the duration of needs improvement status and the previous year's AYP failure help to explain teachers' empowerment. What is interesting is that depending on the nature of the measurement of NCLB, we see both positive and negative associations with empowerment. Consistent with the disorganization perspective, the multivariate analysis demonstrates an improvement of empowerment the longer teachers work under sanctions. However, these data show that teachers feel a decrease in empowerment following each successive failure.

In sum, prolonged needs improvement status is positively associated with teachers' perceptions of the adequacy of their time and sense of empowerment. Conversely, failing AYP in the preceding year has no relationship to teachers' perception of time, but is negatively related to perceptions of empowerment (see Table 2).

## **QUALITATIVE ANALYSIS: TEACHERS' EXPERIENCES – TIME AND EMPOWERMENT INTERTWINED**

The statistical findings presented above suggest a complex web of positive and negative outcomes for teachers' perceptions of time and empowerment. To investigate these results further, I next present analysis of interview data

collected as part of the same research project. Semi-structured interviews were conducted with teachers, principals, and district administrators during the fall of 2006. These interviews were designed to collect data on teachers work experiences under NCLB across three distinct school settings within a single Georgia school district. To highlight variation in NCLB contexts, each of these three schools are at different levels of needs improvement. All teachers in each school were notified of the study and given the opportunity to participate. To maximize variation in teacher perspectives, the original sample included comparable number of respondents from all theoretically relevant categories: grade level, years of experience, and subject area taught. The interview protocol raised many related topics including the impact of high-stakes accountability policy on school working conditions, the strategies used by teachers and school administrators to respond to the policy, and the impact the policy is having on their sense of job satisfaction and career plans. For this chapter, I conducted a systematic analysis of discussions of working conditions found in a subsample of the teacher interviews ( $n = 30$ , 10 teachers per school).<sup>9</sup>

The first school, White Plains Middle School, has never been on the needs improvement list.<sup>10</sup> White Plains is a relatively new middle school that has continually met AYP standards and is labeled a distinguished school. As a result of its distinguished status, White Plains accepts students from other area middle schools that are on the needs improvement list. The second school, Parkside Middle School, is a newcomer to the needs improvement list. At the time of data collection, Parkside had just begun implementing school choice, one of the first sanctions required of needs improvement schools. Greenway Middle School, the final school chosen for study, was in advanced needs improvement standing. At the time of data collection, Greenway was undergoing school restructuring and being monitored by the state. By selecting schools in different needs improvement levels, I highlight variation in NCLB sanctioning experiences.

Each interview focused on how NCLB requirements affect teachers' experiences of time and empowerment in their school setting. Since experiences are likely to vary according to organizational position, I drew a purposive subsample of teachers across subject, grade, and teaching experience. In addition, because special education teachers are often heavily affected by NCLB implementation, I included one special education teacher from each school. When multiple teachers were available from the same categories for selection (subject, grade, teaching experience), respondents were selected at random. In these rich conversations, teachers in each of the

school contexts revealed how variation in the NCLB needs improvement and sanctioning levels impacts their experiences of time and empowerment.

### *Time Use – Teacher Reports*

Teachers in all three schools report a change in the quality of their teaching time as a result of NCLB and associated high-stakes testing. With the increasing focus on standards, testing, and accountability codified by NCLB, all teachers feel that their time is at a premium. The expectation that they will be able to cover the entire mandated curriculum prior to the Criterion-Referenced Competency Test (CRCT), the standardized test used to assess AYP in Georgia middle schools, leaves them feeling rushed. This time crunch not only affects their ability to get curriculum covered at an adequate pace, but also erodes their sense of classroom empowerment. Teachers experience a lack of sufficient time, generating the feeling that they are not concentrating on adequately delivering the required curriculum or meeting their students' needs. Rather, teachers are just "racing against the clock." Thus, time and empowerment are intertwined in their everyday work experiences, regardless of the school needs improvement status.

Elizabeth, an outspoken teacher working at White Plains, expressed the intersection of time and classroom empowerment in this way:

The race to the CRCT score, do you have time for that? And that is what you end up saying to yourself. Oh, it would be nice if I could get John there [to read on grade level], but I don't have time ... I got to go. I've got to go. And that is dehumanizing for me.

Elizabeth feels "dehumanized" because her attention has been pulled away from her students' academic needs and refocused on keeping up with the pace of the curriculum. It is a common sentiment across schools that teachers do not have enough time to "just teach." With this shortage of time comes a loss of control, or empowerment, in their classroom.

As teachers work under NCLB sanctions, this sense of time pressure within their classroom intensifies. Sadie, a science teacher at Parkside remarked:

There is a great deal of pressure about the CRCT, and even though you try not to let it bother you, I think it bothers you all year long. The pressure [is] you want to teach for learning and you don't want to teach for tests, but the reality is you have to teach a little bit for the test because, the test is the test!

The loss of empowerment here takes the form of "teaching to the test" rather than "teaching for learning." In general, the idea of gearing one's teaching towards the CRCT makes teachers uncomfortable and strips them

of their ability to exercise autonomy in curriculum and lesson choices. There is a persistent concern about the CRCT and this seems to pervade many aspects of the school. Mary Beth, an experienced teacher at Parkside, expressed it this way, “It’s not like every day, every minute, but it just kind of hangs. It kind of looms.” Another teacher described the experience of persistent pressure of CRCT looming over their classroom teaching as the “elephant in the room.” The crunch on time resulting from CRCT not only impacts how teachers use their time but also how they feel about their empowerment in the classroom. As the consequences of student performance on the CRCT grow in importance for schools, the “elephant” grows larger in the room.

Teachers at Greenway also experience this intersection of time and empowerment. However, the pressure of working under needs improvement conditions for such a long time has caused teachers at Greenway to develop strategic responses to this shared experience. Teachers identify several strategies that they have devised for dividing and using their time within the pressures that all teachers are feeling to varying degrees. Two strategies dominate teachers’ experience. First, teachers divide up their classroom time in an attempt to meet the diverse needs of their students while moving at a rapid pace. This strategy has special significance for educators in needs improvement schools, since student diversity is directly involved with the reporting of test scores and AYP. Second, teachers eliminate all untested curriculum to focus their limited time on topics that will improve test scores. These strategies help teachers manage the pressures and demands of working in a needs improvement school. Significantly, these strategies are consistent with the disorganization perspective. This may indicate that current accountability initiatives are effectively achieving their goals through increased regulation of teachers’ time.

Jill, a Math teacher at Greenway, describes how she spends a lot of energy subdividing her time to meet the needs of students in various subcategories (e.g. race, ethnicity, class, English proficiency, etc.):

It’s very hard ... I have one student that meets [the reporting requirements] in three categories, and it’s even harder to make sure ... that while I’m trying to help him I have four or five other kids that are in that same class that are considered our bubble kids, and I’m trying to make sure that I go around and make sure those kids get it. But, [you are] just more or less dividing your time up and dividing yourself up into making sure that the others don’t start falling behind.

Jill “divides” herself up to focus her limited time on students that have a larger impact on the schools AYP performance. Testing results are reported

by demographic subgroups. Since AYP is determined by performance in each subgroup, teachers and schools pay special attention to the success of students who are counted in multiple categories.

Many schools have identified “bubble students” (Dworkin, 2005). These students’ test scores fall just below or above the passing grade. At Greenway, bubble students are formally identified by the administration and teachers are assigned to serve as mentors for these students. Teachers in this setting orient the subdivision of their time in a way that responds directly to the school’s needs improvement status.

The intersection of time and empowerment also includes larger school needs at Greenway. Bianchi, another experienced teacher at Greenway, puts it this way:

Basically, we’re racing against a clock. And the clock is the CRCT test. We’re trying to fill them, which is totally educationally inappropriate, to fill them before that test. So you want to fill them as efficiently as you can before that test and they absorb [curriculum] in different ways. In other words, you have different abilities, different learning styles, you know, if it’s a language problem, then obviously verbal is not going to be the best learning style, you know? ... I’m just saying that the diversity makes it more difficult because you have to spend more time planning [and] coming up with more diverse ways of approaching material. Most importantly, you have to spend more time on the material ... because some of them have gaps in their learning already, you’re having to fill in those gaps. So it’s all about time. You know, given enough time we could get them there.

Bianchi notes that teaching a diverse student body requires additional time and special teaching techniques. However, she feels the constraints on her time constrict her ability to fully develop these teaching methods. Instead, Bianchi asserts that she must subdivide her time to meet the educational needs of her diverse student body as quickly as she can. This results in trying to “fill” students before the test, a process that Bianchi deems “educationally inappropriate.” Again, a lack of sufficient time is intertwined with a loss of empowerment in this strategic response exercised by teachers at Greenway.

A second strategy used by teachers working under sanctioning is to focus scarce time on the required curriculum. While all schools feel a similar time crunch and must move quickly through the curriculum, both Parkside and Greenway teachers connect their needs improvement status with a greater concern over the loss of “creativity.” By eliminating all non-mandated curriculum, teachers free up time to deliver the state required and tested standards to their students. Courtney, an experienced teacher at Parkside, summarized her feelings about the loss of curricular freedom like this,



“It restricts much of the creativity that makes learning so much fun. We’ve got to cover this, we’ve got to cover this, we’ve got to cover this because that CRCT is coming up.” Creativity is not always defined by teachers as a special project or lesson, but also any curriculum beyond the basics. Steven, a science teacher at Greenway, states:

It’s not only lesson plan-wise, because you have to be sure all your lesson plans not only are correlating directly with the standards, but you actually ... lose some of your enrichment that you can’t do. I don’t mean that in a totally negative way because what happens is you have to review so much for standards that your time frame doesn’t allow you to do all the extra projects and the extra things that might have taken a little bit longer that really could have made the kid feel like a scientist ... Like run an experiment all the way through. Maybe you just don’t have the time for that. You have to go back to the basics and ... get everybody through.

Eliminating non-mandated lessons and curriculum does further structure and regulate the available classroom time. This strategy focuses the use of classroom time on the material deemed most important by the state. However, it also compromises teachers’ sense of control in their classrooms. This additional curriculum often allows teachers to draw on their own content expertise and flex their professional creativity in the classroom. This strategy to focus their classroom time contributes to a loss of personal empowerment. Camille, a new teacher at Parkside, eloquently expresses this experience:

You’ve got 72 minutes in the day after you’ve done all the attendance and all that silly business and fire drills and all that kind of business ... I mean, what can you do when you shave that down and you divide that up among that many kids, you know? I don’t know a teacher in this school who wants a child left behind ... Not only do you not want to leave them behind, you want them to soar. You are going in many directions at once and it’s just very frustrating. I don’t think it’s frustration that you’re not doing what is expected of you. The frustration is not being able to do what you want to do with the kids.

As she works within the confines of the time in the school day, she has to make tough choices about what curriculum gets taught and what gets omitted. While making these difficult choices, Camille gets frustrated at the loss of her ability to do what she would like to do with the students. This is a loss of empowerment.

The initiative shown by teachers in sanctioned schools to devise ways of responding to the time crunch illuminates one of the interesting statistical findings. The regression results showed that the longer the duration of needs improvement status, the more positively teachers’ assess their time. The interview data reveal that in an attempt to respond productively to the pressures of NCLB requirements, teachers have developed strategies to more highly regulate and structure their time. Thus, while time is short and

the intensity of this time crunch is arguably greater at Greenway and Parkside, their response has resulted in a greater structuring of the available time. It is this agency born out of necessity that helps us explain the positive association between years in needs improvement and time in the survey data. From this perspective, the process of responding to NCLB requirements appears to have effectively embedded itself in classroom processes and focused the use of time on mandated curriculum. As mentioned above, however, in this context there appears to be a trade off between accountability requirements and teacher empowerment. This issue will be further explored by examining teachers' descriptions of empowerment below.

### *Empowerment – Teacher Reports*

The analysis of survey data also shows that longer durations of needs improvement status are linked to more positive assessments of empowerment by teachers. Conversely, the regression analyses demonstrate that teachers experience less empowerment following a failure to meet AYP in the previous year. To help explain these complex findings, we turn again to the interview data.

A close inspection of teachers' descriptions of empowerment reveals the importance of two key aspects of empowerment that help to explain these findings: collaboration and involvement in the school improvement process. Similar to the consistent impact of time on classrooms, the interviews uncover that NCLB requirements have stimulated collaboration across all three schools. Teachers have turned to one another in an effort to understand the state curriculum standards, plan lessons aligned with these standards, and design classroom assessments to test student comprehension. While the quality and results of collaboration vary across schools, all forms seem to bolster teachers' empowerment by creating a "family" or "collective" mentality. At White Plains, the distinguished school not currently experiencing sanctions, teachers describe collaborative experiences as friendly and productive. Teachers report that there is a "family atmosphere" and that collaboration with colleagues is "helpful."

In schools under sanctions, this collaboration takes on new meaning. Mary Beth, at Parkside, has this to say about the nature of collaboration in her demographically diverse and sanctioned school:

I feel like schools with more of an all-white population, the teachers there are more autonomous. They do their own thing. They're more individuals that just operate. But at Parkside, we have to work together like that or we simply won't make it.

For Mary Beth at Parkside, collaboration is not about making the working experience more pleasant but is a method of survival. Teachers in this setting perceive restrictions on their autonomy resulting from the additional effort required to meet the needs of their diverse student population and the expectations codified by NCLB. In these contexts, collaborative work with colleagues becomes a survival strategy.

Likewise, collaboration at Greenway is strategic. When sanctions have increased to such a level that schools are undergoing state contract monitoring and facing school restructuring, the school improvement plan and school improvement team plays a more significant role in teachers' daily experiences.<sup>11</sup> This element of being a needs improvement school becomes a vital part of teachers' collaborative experience. A surface level examination of the number of times teachers mentioned school improvement teams or plans indicates the primacy of this relationship. Out of ten teachers interviewed at each school, only one spoke of school improvement plans at White Plains, whereas this came up in conversation with five teachers at Parkside, the school implementing choice, and eight at Greenway, the school under legal contract with the state. Delving deeper into teachers' descriptions of these experiences, their familiarity and participation in the process confirms that collaborative efforts at school improvement are a vital component of working at these needs improvement schools. In this way, being a needs improvement school generates greater collective response through collaboration in the school improvement process.

Jill explains that teachers compose the majority of the school improvement team at Greenway. "We actually developed a school improvement team. It's a committee of about sixteen teachers including administrators that helped develop the school improvement plan." Stephanie, a new teacher at Greenway, states that all teachers are involved in giving input on the school improvement plan:

I actually got a copy of the school improvement plan today ... A lot of it they're focused on this is what we're doing to improve, these are some of the things we're kind of thinking about doing, getting our reaction so it is a school-wide plan and not just a few people coming up with some ideas to improve. It's everybody, so I have heard a good bit about it.

Just like Stephanie, many of the teachers at Greenway are heavily invested in the collaborative efforts to improve test scores and respond to the demands of NCLB. This involvement in actively working towards getting off the needs improvement list yields a sense of pride in many cases.

Jane describes how her participation on the school improvement team has been a source of great excitement for her:

It's [participation on the school improvement team] been pretty good. In fact, it's [the school improvement plan] all been approved! You know, we do not have to redo anything on it, so we're kind of excited about that. And then it goes before the state school board. I think they said next week or the week after. But, it went rather smooth. We were divided up. I was in part of the students with disabilities section, so we looked at things that we could do for language arts and math. But it was really good. We've been meeting every Wednesday after school. But it's been good. It's a good group of teachers, a really good group, and it just went great so far!

The engagement of teachers in collaborative efforts to improve school performance seems to produce a sense of collective purpose. Even under intense conditions, this strategic collective response contributes to a sense of control over the school environment. Teachers are actively involved in the school improvement process and this is important to them.

This collective response sheds some light on the statistical findings concerning empowerment. First, the greater sense of empowerment echoes the increasing collaborative efforts evident in schools responding to sanctions. Second, an investment in school improvement actually makes the failure of AYP more meaningful. When teachers' investment in school improvement does not improve AYP outcomes, teachers feel a dramatic reduction in empowerment in these "failing" schools.<sup>12</sup> For example, Chanda, a science teacher at Parkside, explains that she feels a loss of empowerment following each successive AYP failure. Chanda states:

I think what bothers me a little bit when we fail AYP is [that] ... what you have to focus on shifts. Last year [when] the CRCT scores came in we didn't pass it because of language arts or reading. So last year the curriculum had to be focused on reading, reading, reading. Well, last year's spring CRCT, it was math, so now we have to work math into the curriculum, well what about the reading? ... So that's the one thing about AYP that bothers me a little bit is that whatever we failed, that's what we have to work on in all the classes ... I think the focus depending on what we failed that year kind of bothers me ... It's not the way I teach, but it's more like what I have to focus on.

Chanda identifies this change in empowerment as partially the result of how the whole school reflexively and automatically shifts academic focus. Thus, the need to respond collectively to the policy can produce positive outcomes by bringing teachers together, but it can also result in the loss of personal empowerment in the classroom. When teachers shift the focus of their teaching in response to the collective need to respond to AYP, they lose some of their autonomy and sense of professional expertise in the classroom.

Similarly, Jill, a teacher at Greenway, feels a dramatic loss of empowerment when, after a huge investment of time and energy into school improvement, her school still fails to meet their AYP targets. She remembers the first faculty meeting where her principal explained the previous year's testing results:

It was very overwhelming. It was very overwhelming because I'm thinking, we work as hard as we possibly can. I was thinking, you know, last year was extremely stressful. This year has been even more stressful since this is our actually our seventh year [in needs improvement], and it sometimes makes you feel very down looking at those students [scores] and thinking, you know, "Three more [students]. Only three more." And I know that seems like that isn't many but, at the same time, it is. At the end of the meeting, he [the principal] actually called out the students that did not make it last year ... and each teacher had to stand if they taught one or more of those students and, by the time we finished, almost every teacher had stood. But yet those are the kids that did not make it. And that makes you feel like, you know, each one of us had seen that student, but at the same time, "What am I doing wrong? You know, why is that child not making it?" And that's where it's kind of depressing.

This vivid collective experience encapsulates the power of being labeled a "failing" school. The unsuccessful collaborative effort at improving test scores produces a striking loss of empowerment. Thus, a school's collective response to needs improvement status and improvement efforts described by teachers may assist in explaining the positive association between needs improvement and empowerment found in the survey data. However, the negative statistical relationship between AYP and empowerment may be capturing the dramatic loss of empowerment following AYP failure (Table 3).

## DISCUSSION

Overall, I found that top-down accountability policies have produced positive and negative changes in teachers' perceptions of their working conditions. The nature of these changes, and teachers' responses to them, vary systematically across schools depending on their students' performance and severity of sanctions. Teachers report that standardized curriculum, high-stakes testing, and the pressure to meet AYP targets limit the meaningful use and availability of time. However, in schools under stronger sanctions, educators generate specific strategies to respond to the novel demands placed on their time. Two strategies emerge in needs improvement schools. First, the need to adequately cover curriculum prior to CRCT

**Table 3. Summary of Findings.**

	Needs Improvement	AYP Status
<b>Time</b>	<p><i>Interviews</i></p> <p>White Plains (Distinguished School)                      Limited time to cover curriculum before CRCT test                      Time and Empowerment Intertwined                      Parkside (Implementing School Choice)                      Limited time to cover curriculum before CRCT test                      Time and Empowerment Intertwined                      STRATEGIES: Eliminate non-mandated curriculum                      Greenway (Implementing Contract Monitoring and School Restructuring)                      Limited time to cover curriculum before CRCT test                      Time and Empowerment Intertwined                      STRATEGIES: Eliminate non-mandated curriculum                      Subdivide limited time to meet diverse student needs</p>	<p>QLTE Survey: Positive &amp; Significant</p> <p>QLTE Survey: Negative &amp; Non-significant</p>
<b>Empowerment</b>	<p><i>Interviews</i></p> <p>White Plains (Distinguished School)                      Demands placed on schools has stimulated collaboration                      Quality of Collaboration: "Helpful"                      Parkside (Implementing School Choice)                      Demands placed on schools has stimulated collaboration                      Quality of Collaboration: Necessary for "Survival"                      Moderate Involvement in School Improvement                      Greenway (Implementing Contract Monitoring and School Restructuring)                      Demands placed on schools has stimulated collaboration                      Quality of Collaboration: Explicitly integrated into School Improvement, Necessary to get off the needs improvement list                      Intense Involvement in School Improvement</p>	<p>QLTE Survey: Positive &amp; Significant</p> <p>QLTE Survey: Negative &amp; Significant</p>

causes teachers to divide up their time and eliminate non-mandated curriculum. This finding suggests that the needs improvement status and accompanying sanctions produce important changes in how teachers manage their time. Thus, the methods do appear to make teachers more accountable to the state designed standards and testing.

Simultaneously, the nature of empowerment in the school is changing. Accountability requirements stimulate educators to adopt a more collective empowerment stance; however, this collaboration is experienced differently in needs improvement contexts. The school context transforms collaborative efforts into a survival strategy to successfully navigate the school improvement process and succeed at meeting AYP. These data suggest that growth in collaboration over time is another positive outcome of these accountability initiatives.

Ironically, the adaptive strategies developed by teachers and schools also seem to produce unintended negative consequences. Strictly regulating time and eliminating non-mandated curriculum can reduce teachers' sense of classroom empowerment. Similarly, strengthening group identification through increased collective empowerment makes the disappointment of AYP failure even greater. Looking at the complete picture revealed in the combination of quantitative and qualitative findings shows that NCLB's accountability initiatives produce both positive and negative outcomes for teachers.

These findings highlight the importance of understanding the educator, administrator, and policy maker perspectives and actions in light of their organizational position. As this study illustrates, these views are not necessarily competing, but rather simultaneous and distinct views of the same events. The educator, administrator, and policy maker subscribe to the same goals: increased achievement and greater equality. However, one's position in the organization generates a distinct perspective, which results in a different set of mechanisms to achieve these goals.

These contrasting viewpoints are likely a product of the highly decentralized American educational system. If the United States was characterized by a centralized and regulated structure, teachers might not perceive attempts to regulate class time or curricular choices as problematic. Given that the U.S. system is decentralized, it is logical to find distinct processes that represent each perspective. For example, consistent with the disorganization viewpoint, one would expect to find evidence of attempts to centralize and regulate school processes. This perspective was represented in the data. Teachers do report further structuring of their time and the increased need to collaborate to meet policy requirements. This approach to

school reform, and the actions it yields, has a significant and beneficial impact on teachers' working experiences.

From the disempowerment viewpoint, one would expect to find evidence that teachers perceive these efforts to regulate their work as problematic intrusions into their classrooms. We also saw evidence consistent with this perspective. Teachers in this sample describe how the restructuring of their time results in a reduction of autonomy in curricular choices, and thus a loss of empowerment. Unlike the disorganization view, teachers perceive this regulation as a barrier to their ability to effectively meet diverse students' needs. In addition, teachers indicate that collaborative efforts directed at the school improvement process can produce a significant reduction in empowerment if they are unsuccessful.

Most of the research on NCLB has emphasized outcomes for students, which is appropriate given the aims of the policies. The goal of this chapter was different: it has examined the impact of these policy requirements on teachers. In some ways, the policy is having the intended and beneficial effects – it is structuring the use of teachers' time and promoting collaboration. But in other ways, the policy also generates unintended impacts on teachers' working conditions. In particular, the regulation of time and the elimination of non-mandated curriculum reduce teachers' sense of classroom empowerment. Similarly, bolstering group identification through increased collective empowerment makes the sting of AYP failure even greater for teachers. These findings raise several important questions for continued research on this topic. First, what are the long-term ramifications of changing work conditions for teachers' career trajectories? As schools continue to grapple with issues of teacher attrition, we must consider how working conditions may affect teachers' decisions to stay in or leave the profession. This line of research may prove beneficial for improving retention of teachers, especially in struggling schools.

In addition, by focusing on how policy requirements impact teachers, this study does not investigate how teachers in turn influence the ultimate goals of education: equality and achievement. Further research should be done to explore the ways in which teachers' experiences serve as intervening variables between policy initiatives and student achievement and equality. Specifically, how are student outcomes directly affected by the reorganization of teachers' time and the focus on mandated curriculum? Are there indirect effects on student outcomes through teacher attrition or out-of-field teaching? These issues, and a host of others, need to be examined to develop a comprehensive picture of the multiple and varied influences of accountability policy.



## NOTES

1. Throughout this chapter, I use teachers' perceptions to gauge changes in their work environment. Teachers, like other individuals, behave on the basis of their perceptions or the meanings they create around these perceptions (Howard & Renfrow, 2003; Heider, 1958; Stryker & Vryan, 2003). For teachers, their perceptions of the working conditions in schools have real consequences for their job satisfaction and career plans (Ingersoll, 2001). Thus, it is crucial to consider the ways in which accountability policy reshapes teachers' perceptions and assessments of working conditions.

2. While the debate continues over the effects of schools, there is a long history of research which demonstrates that many of the causes of educational inequalities lie outside of the school system (Coleman et al., 1966; Jencks et al., 1972; Entwistle & Alexander, 1992, 1995; Gamoran, 1995). However, most education reform, including NCLB, focuses on schools and teachers. This is a reasonable response by educational policy makers since their realm of control is limited to this particular institution. The efforts to resolve educational inequalities originating or reinforced within schools are extremely important, yet their effectiveness is hampered by the strength of broader societal inequalities. Ironically, in more centralized and highly regulated systems, the relative power of schools to resolve educational inequalities increases (e.g. Park, 2006; Heyneman & Loxley, 1983). Thus, in the U.S. context, our educational structure may render attempts at rectifying educational inequalities by restructuring only educational practices and processes less effective. Regardless, such efforts are a necessary and important part of the ongoing attempts to reform U.S. education.

3. The QLTE survey is part of a larger movement across the United States to collect data on school working conditions. I would like to thank the Board of Regents of the University System of Georgia and the Office of P-16 Initiatives for granting me access to this valuable data. This Initiative is funded by BellSouth Georgia and the BellSouth Foundation. More information about the initiative can be found at <http://www.qlte.org/>.

4. I conducted principal components analysis with varimax rotation on all items included in the survey falling under each working condition domain. Not all items measured the domain equally well, thus PCA was used for data reduction purposes. A complete listing of the questions included in the time and empowerment indices are available in Table 1.

5. Sanctions are outlined by NCLB; however, states retain some power in deciphering the actual services to be provided to schools. While this variation across states is significant for national assessments of NCLB, this study focuses on a single state. Since all data was collected within the state of Georgia, sanctions are consistent across schools in similar needs improvement status.

6. Needs Improvement and AYP are related to one another ( $r = 0.48$ ). A school's needs improvement status is the total of all AYP results beginning after the second consecutive failure. In other words, if a school does not meet AYP for several consecutive years, then they are, by definition, further along in needs improvement. What is unique about these measures of NCLB is that they capture distinct segments of time. Number of years in needs improvement captures the

duration of sanctions in the school. AYP just records the result of the previous year's effort. Since each measure tests a different type of relationship with the dependent variable, both measures were utilized. The models presented here include both needs improvement and AYP. Separate models were tested to see if each measure made a distinct contribution to time and empowerment when considered alone. In the models for time, the results were unchanged. However, in the models for empowerment, years in needs improvement no longer produced a significant effect. This difference in results is the product of the moderate correlation between the variables and the direction of the coefficient. Including both variables in the same model presents a more accurate assessment of the simultaneous effect of the NCLB conditions. Thus, I chose to present those models.

7. In analyses not presented here, I included a control variable for district to assess the impact of district on working conditions. In all models, the district had no effect on teachers' reports of time and empowerment. In the interest of parsimony the variable was excluded. It would also be beneficial to include measures for per pupil expenditures and teacher salaries. Unfortunately, per pupil expenditures is not available at the school-level and teacher salaries were not collected in the original survey instrument. However, since all teachers included in the sample work in Georgia they share a similar pay scale. In addition, differences in the pay scale are partly captured by the inclusion of years of teaching experience.

8. This finding may be consistent with the relative status expectation argument. This argument suggests that due to persistent patterns of institutionalized discrimination, males and whites are typically found working in positions where they have higher levels of power and control than females and non-whites, respectively. This pattern of increased power and control due to privileged status characteristics produces the expectation of control. Thus, the relative lack of power and control resulting from the organizational structure is particularly salient to their work experience and yields a perception of lesser control in those contexts (Gordon, 2006).

9. Beginning with an initial start list of codes, I began the iterative process of examining and coding the data. As I proceeded, I generated analytic memos and continually refined my code list as new codes emerged from the data. Throughout the coding and analysis process I created matrix and network displays to reduce the data and identify key themes, ideas, and processes. I used MaxQDA to facilitate data management, coding, and analysis (Glaser, 1978; Miles & Huberman, 1994; Lincoln & Guba, 1985).

10. I use pseudonyms for all school and teacher names to protect confidentiality.

11. School improvement is the phrase used to describe the process of responding to the needs improvement label. In Georgia, schools in needs improvement must formulate a school improvement plan. Typically, the school improvement plan is generated by a school improvement team.

12. While policy and state reports do not use the phrase "failing school" to describe schools that did not meet AYP, the popular media, teachers, and community members do use this language. Given the potency of its connotation, I feel this phrase is important to use here in the discussion of teacher empowerment.

## ACKNOWLEDGMENTS

I gratefully acknowledge the support of the National Science Foundation as well as the Emory University Graduate School of Arts and Sciences. I also appreciate assistance from the Board of Regents of the University System of Georgia and the Office of P-16 Initiatives for granting me access to the Quality Learning and Teaching Environments data and the numerous teachers, school, and district administrators that gave their valuable time to speak with me on these important topics. Lastly, I am grateful for the tremendous help and guidance given to me by my dissertation committee members: Richard Rubinson, Monica Worline, Regina Werum, and Matthew Archibald.

## REFERENCES

- Baker, D. K., & LeTendre, G. K. (2005). *National differences; global similarities: World culture and the future of schooling*. Stanford, CA: Stanford University Press.
- Bidwell, C. (1965). The school as a formal organization. In: J. G. March (Ed.), *Handbook of organizations* (pp. 972–1018). Chicago: Rand McNally.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, DC: Government Printing Office.
- Costigan, A. T., & Crocco, M. S. (2004). *Learning to teach in an age of accountability*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Dreeben, R. (1973). The school as a workplace. In: R. M. W. Travers (Ed.), *Second handbook of research on teaching* (pp. 450–473). Chicago: Rand McNally and Company.
- Dworkin, G. A. (2005). The no child left behind act: Accountability, high-stakes testing, and roles for sociologists. *Sociology of Education*, 78, 170–174.
- Elmore, R. (2000). *Building a new structure for school leadership*. Washington, DC: The Albert Shanker Institute.
- Entwistle, D., & Alexander, K. (1992). Summer setback: Race, poverty, school composition, and mathematics achievement in the first two years of school. *American Sociological Review*, 57, 72–84.
- Entwistle, D., & Alexander, K. (1995). A parent's economic shadow: Family structure versus family resources as influences on early school achievement. *Journal of Marriage and Family*, 57, 399–409.
- Fuller, B., & Rubinson, R. (1992). *The political construction of education: The state, school expansion, and economic change*. New York: Praeger.
- Gamoran, A. (1995). Effects of schooling on children and families. In: A. Booth & J. Dunn (Eds), *Family-School links: How do they affect educational outcomes?* (pp. 107–114). Lawrence Erlbaum.

- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.
- Gordon, K. (2006). *Relative Expectations: School Composition, Context, and Teachers' Perceptions of Control*. Unpublished paper presented at the annual meeting of the Southern Sociological Society, August, New Orleans, LA.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Heyneman, S. P., & Loxley, W. A. (1983). The effect of primary-school quality on academic achievement across twenty-nine high- and low-income countries. *The American Journal of Sociology*, 88(6), 1162–1194.
- Hirsch, E., Emerick, S., Church K., & Fuller, E. (2006). Teacher Working Conditions are Student Learning Conditions available online from The Southeast Center for teacher Quality at <http://www.teachingquality.org>
- Hoffer, T. B. (2000). Accountability in education. In: M. Hallinan (Ed.), *Handbook in the sociology of education*. New York: Kluwer Academic.
- Howard, J., & Renfrow, D. G. (2003). Social cognition. In: J. Delamater (Ed.), *Handbook of social psychology* (pp. 259–281). New York: Kluwer Academic.
- Ingersoll, R. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal*, 38, 499–534.
- Ingersoll, R. (2003). *Who controls teachers' work*. Cambridge, MA: Harvard University Press.
- Ingersoll, R. (2005). The anomaly of educational organizations and the study of organizational control. In: L. V. Hedges & B. Schneider (Eds), *The social organization of schooling*. New York: Russell Sage Foundation.
- Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., & Mickelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books Inc.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Lortie, D. (1975). *School teacher*. Chicago: The University of Chicago Press.
- Meyer, J., & Rowan, B. (1978). The structure of educational organizations. In: M. Meyer (Ed.), *Environments and organizations* (pp. 78–110). San Francisco: Jossey-Bass.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- National Center for Education Statistics. (1997). *Teacher professionalization and teacher commitment: A multilevel analysis*, February. Washington, DC: U.S. Department of Education.
- No Child Left Behind Act 2001. (2002). 20 USC § 6301.
- OECD – Organization for Economic Cooperation and Development. (2003). *Education at a glance: OECD indicators*. Paris, France: OECD.
- Park, H. (2006). Cross-national variation in the effects of family background and schools on student achievement: The relevance of institutional and policy contexts. Doctoral Dissertation, University of Wisconsin, WI; Dissertation Abstracts International, 66(8), pp. 3098–3297.
- Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education*, 16, 353–389.
- Schaub, M., & Baker, D. P. (1991). Solving the math problem: Exploring mathematics achievement in Japanese and American middle schools. *American Journal of Education*, 99, 623–642.
- Stryker, S., & Vryan, K. D. (2003). Symbolic interactionst frame. In: J. Delamater (Ed.), *Handbook of social psychology* (pp. 3–28). New York: Kluwer Academic.

- Sunderman, G. L., Kim, J. S., & Orfield, G. (2005). *NCLB meets school realities: Lessons from the field*. Thousand Oaks, California: Corwin Press, A Sage Publications Company.
- Waller, W. (1932). *The sociology of teaching*. New York: Russell and Russell.
- Weber, M. (1946). The 'rationalization' of education and training. In: R. Arum & I. Beattie (Eds), *The structure of schooling* (pp. 16–18). Boston: McGraw Hill.

## CHAPTER 5

# RAISING ACHIEVEMENT OR CLOSING GAPS? IDENTIFYING EFFECTIVE ACCOUNTABILITY TOOLS

Melissa K. Henne and Heeju Jang

### ABSTRACT

*California enacted a standards-based accountability regime in 1999, aiming to boost achievement overall and narrow gaps among subgroups. Yet we know little about the efficacy of specific accountability practices and reform tools observed by teachers and principals. The loose-coupling critique of school organizations, positing that local educators steadily buffer interventions mounted by state actors, is challenged by a selective-coupling representation where school-level actors do experience rules and incentives that encourages compliance with state advanced curricular standards, pedagogical practices, and standardized testing. After surveying educators across a band of similar elementary schools, we can account for sizeable shares of the variance in mean Academic Performance Index (API) scores among schools and in the size of achievement gaps within schools. We found that achievement levels are higher when principals report a stronger district focus on a unified curriculum and their teachers share high expectations for learning. Gaps between white and Latino students are smaller when teachers report steady attention to meeting accountability*

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**

**Research in Sociology of Education, Volume 16, 133–155**

**Copyright © 2008 by Emerald Group Publishing Limited**

**All rights of reproduction in any form reserved**

**ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16005-0**

*targets. Latino achievement is more sensitive to these accountability practices, compared with the performance of white students. Even after sampling schools with similar student populations, the social-class background of students continued to heavily influence achievement levels, explaining greater shares of the variance than accountability practices.*

Activists and scholars – over the past half-century – have struggled to pinpoint efficacious school practices that can be replicated across diverse schools and students. The original architects of the *systemic reform* model of accountability aimed to incorporate core findings from the effective-schools (ES) literature, accenting the importance of a coherent set of learning objectives, schoolwide collaboration among teachers to ensure alignment, and stronger incentives for sound pedagogical practices (e.g., [Smith & O’Day, 1991](#)). But as states in the U.S. and other nations advance accountability policies, few studies have identified the specific practices or policy tools felt within schools that raise student achievement or narrow gaps among subgroups.

We delve into this question in the wake of California’s own accountability effort begun in 1999. We know that the work of school principals and the classroom practices of teachers has, on average, shifted to align more carefully with the state’s curricular standards, new materials married to the standards, transparent information about student performance, and rewards or sanctions for schools contingent on learning curves (see [Hamilton et al., Chapter 2](#)). What’s not well understood is how the changing practices of principals and teachers – operating between state capitals and classrooms – are empirically related to achievement patterns.

We first review the ES literature, emphasizing that it matured as states were beginning to confront the loosely coupled layers of the school institution, going back to the 1980s in the U.S. and Britain (see [Fuller, Chapter 1](#)). This literature illuminated coherent organizational practices and engaged staff in many successful schools but yielded less knowledge about whether these practices could be scaled-up to become systemic reforms. Little work was conducted to identify district- or school-level practices that narrow gaps among student subgroups, a major focus of contemporary accountability initiatives. ES scholars, estimating achievement across large numbers of schools, attempted to take into account observable facets of pupils’ social-class background via conventional controls within multivariate regression designs. But estimates of “best practices” are infamously biased when school quality indicators are collinear with student attributes.

We emphasize theoretically that the rise of state-led accountability measures – as well as policy regimes exercised by central governments, such as No Child Left Behind – present more penetrating organizational changes that may render as obsolete the earlier loose-coupling representation of the school institution. As governments press for stronger technical efficiency inside classrooms, we may see tighter selective couplings between policies (e.g. curricular standards or instructional programs) and teacher responses to rules or controls over their everyday work (Meyer & Rowan, 2006; Fuller with Bridges & Pai, 2007). This causal sequence, when experienced by local educators, may result in a variety of effects, from changing the collaborative relations among teachers to alienating them from a more didactic form of education. But the first step in this causal chain is whether teachers and principals observe organizational change linked to accountability policies and tools.

Finally, we will describe the design of our California-based study – *Similar Students, Differing Results* (SSDR) – where a large sample of elementary schools was drawn randomly from a narrow band of schools which shared similar class characteristics. We attempted to identify principal or teacher-reported practices – be they exercised by district leaders or school-level educators – which may be associated with higher mean levels of achievement and smaller gaps among student subgroups.

Thus in this new environment of standards-based accountability, we aimed to identify effective district- or school-level practices, distinguishing between average achievement levels and gaps in learning, between white and Latino children. We rely on the reports of principals and teachers and compare the efficacy of their reported practices to estimate student performance, net the a priori influence of student background characteristics. We attempt to isolate on effective practices by sampling schools from a narrow band of institutions with similar student populations. We further control on several social-class attributes at the family level to conservatively test for associations between district and school practices and achievement levels.

## **DISSECTING EFFECTIVE SCHOOLS**

### *Efficacious Organizations in Decentralized Environments*

The first wave of ES research aimed to identify organizational practices – largely operating within schools and decades before government-led



accountability efforts took root – which appeared to boost student performance (Levine & Lezotte, 1990; Rosenholtz, 1985). A fledgling second wave of ES work has recently asked whether standards-based accountability policies can advance school or district practices that lift student motivation and learning (Bell, 2001; Carter, 2000; Education Trust, 1999; Izumi with Coburn & Cox, 2002). The first wave certainly informed the design of systemic reforms, including an emphasis on coherent curricular objectives, teacher collaboration, and more transparent assessment of students' growth curves (Purkey & Smith, 1983). The core empirical questions and methods remain similar between the two waves of ES investigations, as both traditions rely on identifying positive outlying schools that display performance that exceeds what one would expect, based on student characteristics. As a result, partialing-out the prior effects of home background remains slippery in non-experimental studies.

Still, this line of work consistently identifies school climate as having a measurable impact on student achievement (Levine & Lezotte, 1990). This includes the fundamental notion that educators in effective schools express high expectations for all students (Purkey & Smith, 1983), and have a collective sense of responsibility to ensure that everyone is involved in improving student achievement (Johnson, 1999). ES scholars highlight the important role of teachers who are deeply committed to helping all students achieve (Cawelti, 1999), placing learning at the center of the school's everyday activities, compared with discipline or simply herding students through the institution (Bell, 2001; Johnson, 1999), and creating a tacit culture of achievement (Public Schools of North Carolina, 2000). ES research continues to emphasize the use of curricular standards to guide teaching and learning, at times reinforced through teacher collaboration, as a key element of an effective organization. Several studies have found that instructional coherence is directly related to standards-based instruction (Education Trust, 1999; Levine & Lezotte, 1990).

Effective schools tend to assess students frequently and in ways that directly aid teachers in adjusting attention given to particular students or subgroups (Izumi with Coburn & Cox, 2002; Virginia Department of Education, 2000). Teacher development focused on aligning instruction to standards, along with sustained teacher collaboration, also appear to boost achievement (Johnson, 1999; Phi Delta Kappan, 2001). And many studies have underscored the principal's potent role when he or she can focus on instructional improvement by spending ample time in classrooms and working with teachers (Carter, 2000; Levine & Lezotte, 1990; Perry & Pelavin, 2003).

*Technical Efficacy in Centralizing Environments*

Previous research argues that public schools can often win strong social and political legitimacy, and in doing so buffer technical demands that would be placed on the school's inner-workings (Meyer & Rowan, 1977). By buffering teachers from outside intervention or demands, principals and district staff could protect the popular credibility that their institutions once enjoyed. Yet the architects of standards-based accountability, informed by this organizational perspective, began to selectively couple central policies with school-level practices in the 1980s. Research on accountability has found that such policies are often linked with noticeable changes in school and classroom-based practices. This presents a fundamental challenge to the loosely-coupled, legitimacy-seeking conception of the school institution that have come to face rather demanding technical environments (Rowan, 2006).

In this light, our study delineated specific tools of California's accountability policies, and then aimed to explore the related variability in their implementation with student achievement levels. More broadly, the centralizing environments in which American schools now operate represent a bundle of theoretical and empirical challenges for contemporary students of organizational behavior. Demands for performance gains have intensified, and technical fixes abound, from new input mixes (like reducing class size) to regulation of pedagogy (like the use of prescribed curriculum packages).

Certain elements of schooling certainly remain taken for granted or institutionalized, from the separation of students by age and grade level, to the increasing atomized conception of knowledge, delivered to student in routinized ways (Rowan, 2006). Thus, tandem questions on this new empirical frontier include: What kinds of observable accountability tools are inconsistently or selectively coupled across district- and school-level actors? And which accountability tools or practices, from district leadership in pressing curricular alignment to steady assessment of student performance, directly affect student achievement? These are questions that arise once the technical demands for regulated change inside schools are recognized and experienced by local educators, shifting our theoretical attention from the institutional environment which once simply required surface-level appearances of rationality or compliance. At the same time, we can not reject the loose-coupling representation of the school institution until we see which particular policy tools are observed by local educators and demand real shifts in their everyday work.

*Persisting Holes in Effective-Schools Research*

Beyond digging into these empirical questions, we aimed to avoid prior weaknesses long evidenced in the ES tradition. Our design aimed to avoid the problem of collinearity between school qualities (accountability practices in our case) and students' social-class background. Both aggregate school effects and the influence of particular ingredients are commonly misestimated. To minimize this risk, we sampled elementary schools within a narrow band of similar schools, as detailed below. We also sought to minimize measurement shortcomings by first interviewing a subgroup of principals and teachers, hoping to triangulate on efficacious district or school practices which stem from state accountability policies. Finally, many states and nations are struggling to narrow achievement gaps between student subgroups. It may be that certain policies or local practices are more effective in narrowing gaps, compared with accountability tools that focus on raising mean levels of achievement.

## **STUDY DESIGN – IDENTIFYING EFFECTIVE ACCOUNTABILITY TOOLS**

Our analysis stems from the California study, SDDR, led by [Williams, Kirst, and Haertel \(2005\)](#).<sup>1</sup> This investigation examined practices and tools linked to accountability reforms, exercised at the district or school level, that were associated with higher student achievement. It was based upon teacher and principal survey responses from 267 elementary schools.

*Analytic Priorities and School Sample*

The SDDR school sample was drawn in fall 2003 from a narrow band of schools falling in the 25th–35th percentile of the state's School Characteristics Index, a composite gauge of the attributes of pupils and their families. This particular band of schools was selected because, while they face substantial student demographic challenges, they are not the most severe in the state ([Williams et al., 2005](#)). By utilizing this truncated range of students and communities, the study team hoped to control for social-class factors. While schools educated a similar student population, their Academic Performance Index (API) scores ranged widely, displaying a difference of

just over 250 points in this scale (determined mainly by standardized test scores) that ranges from 200 to 1,000 points.

Our analysis draws upon four different outcome measures, corresponding to four similar regression models. The models for schools' total API score and the API score for Latino subgroup draw upon a similar sample of schools. The models for the API score for white students were run separately, based on a slightly smaller sample, as less than half of the original school sample included a sufficient count of white students. The model for the achievement gap (API units) between white and Latino students draws from a reduced sample of schools, given that we wanted to estimate the gap in API scores for different ethnic groups within each participating school. Estimated coefficients did not change substantially when all four models were run with a constant-school sample while increasing standard errors and constraining degrees of freedom.<sup>2</sup>

California's Public School Accountability Act (PSAA) requires that all schools report a subgroup score for any "numerically significant" subgroup, meaning a group that has at least 100 or more students or 50 or more students who make up at least 15 percent of all students tested under the state assessment program. Many schools in our sample did not meet the minimum subgroup count for Asian or African-American students.<sup>3</sup> Therefore, we decided to exclude these two ethnic groups from our analysis. As a result, our achievement gap findings are based upon teacher and principal interview responses from elementary schools that reported a white and Latino API subgroup score.

We address our core questions by first examining the mean achievement levels and gaps between white and Latino students. We then use multivariate regression to analyze the association between reported practices linked to accountability policies on mean achievement levels and the size of gaps between the two subgroups. All estimates are modeled at the school level, including teacher reports aggregated to the school level (means). Overall, this analytic strategy allowed to test whether practices – reportedly emanating from the district, principal, or teachers – are related to mean achievement and/or narrower gaps between white and Latino students. We include student background characteristics in regression estimates to further take into account the presumably a priori influence of student's social-class position.

This study has four outcome measures and multivariate models. Four regressions were modeled to predict an overall API scores, API scores for white and Latino subgroups, and the achievement gap (API units) between the two subgroups. The identical set of predictors was used in each model.

The following controls for student background were used in each model: the share of school enrollment that is white or Asian (collinear with black and Latino enrollment shares), the percentage of students who receive free or reduced-price lunches, English learners, migrant students, geographically mobile students, and students whose parents have a high school diploma or a college degree. Since two schools with a majority enrollment of Latino students displayed particularly high API scores, these schools were controlled for outlying fixed effects. We will first report on practices reported by teachers for which reliable indices could be constructed; we then present similar regression models for indices constructed from the principal reports of accountability practices.

### *Predicting Achievement Levels – District and School Practices*

*Teacher and principal reports.* These tandem surveys covered a range of issues related to district- and school-level practices. To arrive at the indices detailed below we conducted principal components analysis to assess how interview items clustered together. Only indicators were retained for the present analysis when they were conceptually related to California's accountability press and when inter-item reliabilities (Cronbach's alpha) were moderate to strong. We were cognizant of the kinds of effective practices stemming from the earlier literature, but we also focused on district or school practices that clearly stemmed from post-1999 state accountability policies. As expected, there is both some overlap and difference between the teacher and principal factors. Finally, in order to ensure that factors were measuring different aspects of student achievement, we ran a correlation between teacher and principal factors. Factors that were highly correlated with each other were dropped to avoid multicollinearity in the regression models.

The first set of regression models reports on the possible association between achievement and several teacher-reported practices.

*Teachers use of curricular standards.* This index measures the extent to which teachers report utilizing state standards to guide their classroom instruction, including the identification and use of core or "key" standards, and consulting with their principal as a knowledgeable source for learning about the standards that should be addressed in one's pedagogy (Cronbach's alpha = 0.60).

*School focuses on achievement.* This index taps the extent to which teachers perceive that staff focus on student achievement, including a

school's mission and specific plans for instructional improvement. It also includes the role of the principal in communicating a clear vision, expressing clear and consistent expectations, and setting high standards for achievement (Cronbach's  $\alpha = 0.88$ ).

*School prioritizes meeting accountability targets.* This composite measure gauges the extent to which school staff push to meet state and federal accountability targets, including a school's awareness of and responsiveness to API and student subgroup growth targets (Cronbach's  $\alpha = 0.86$ ).

*Teachers express high expectations for student learning.* This index measures the extent to which teachers have high expectations for student achievement, as well as whether they take responsibility for and are committed to improving student achievement. Additionally, teachers communicate to students that education is important and discuss assumptions about race and student achievement (Cronbach's  $\alpha = 0.69$ ).

*District leaders prioritize student learning.* This index gauges the extent to which district staff emphasize student learning and gains in achievement, including supporting schools and teachers in boosting student learning and providing student assessment data (Cronbach's  $\alpha = 0.77$ ).

*Teachers modify pedagogical practices.* This composite measures the extent to which teachers are allowed to experiment with their instructional strategies, including opportunities to integrate culturally relevant material and incorporate new pedagogical techniques (Cronbach's  $\alpha = 0.91$ ).

*School staff focus on student subgroups.* This index measures the extent to which teachers and staff prioritize the progress of subgroups, with teachers using assessment data to analyze subgroup achievement, setting measurable goals for subgroups, and receiving professional development around the use of assessment data. Furthermore, there is a schoolwide effort to discuss and monitor subgroup achievement (Cronbach's  $\alpha = 0.87$ ).

*District staff support teachers.* A composite measure that taps the extent to which a district supports their teachers, including praising and recognizing their work. Additionally, the district is aware of school-level problems and challenges, helps build community confidence in their school, and encourages teachers to partake in district decision-making (Cronbach's  $\alpha = 0.89$ ).

*School displays a positive environment.* This variable measures the extent to which school actors reportedly create and maintain a positive and orderly school climate. The school enforces policies regarding student behavior, including rules for a dress code, bullying, and attendance. Another item asked about whether students respect cultural differences among students (Cronbach's  $\alpha = 0.81$ ).

*Teacher development for serving English learners.* This index measures the extent that professional development activities, including assistance for coaches, observing other teachers, and receiving feedback on individual instruction, influences teachers with their instruction for English learners (Cronbach's alpha = 0.80).

The second set of regression models reports on the possible association between achievement and several principal-reported practices.

*School focuses on achievement.* This composite measures the extent to which a principal reports that staff focus on student learning outcomes, including well-defined plans for instruction, ensuring that classroom teaching is guided by state standards (Cronbach's alpha = 0.73).

*District guides school curriculum.* This index measures the extent to which district expectations shape a school's curriculum, ensuring that it is aligned and coherent. Additionally, the district evaluates school leaders and has highly skilled staff (Cronbach's alpha = 0.77).

*Teachers express high expectations for student learning.* This index measures the extent to which teachers have high expectations for student achievement, as well as whether they take responsibility for and are committed to improving student achievement. Additionally, teachers communicate to students that education is important and discuss assumptions about race and student achievement (Cronbach's alpha = 0.79).

*District expects schools to meet accountability targets.* This variable measures the extent to which a principal understands district expectations regarding API and AYP subgroup and growth targets (Cronbach's alpha = 0.98).

*School prioritizes meeting accountability targets.* This composite measure gauges the extent to which school staff push to meet state and federal achievement growth objectives, including a school's awareness of and responsiveness to API and student subgroup growth targets (Cronbach's alpha = 0.77).

*District expects schools to improve.* This index measures the extent to which districts use student achievement results to evaluate school performance. Additionally, districts expect schools to improve and principals to understand these expectations (Cronbach's alpha = 0.62).

*School staff focus on student subgroups.* This index measures the extent to which teachers and staff prioritize the progress of subgroups, with teachers using assessment data to analyze subgroup achievement, setting measurable goals for subgroups, and benefiting from professional development on the use of assessment data. Schoolwide efforts to discuss and monitor subgroup achievement is also tapped by this measure (Cronbach's alpha = 0.88).

*School displays a positive environment.* This variable measures the extent to which a school creates and maintains a positive and orderly school climate. The school enforces policies regarding student behavior, including rules for dress code, bullying, and attendance. Furthermore, students respect cultural differences among students (Cronbach’s alpha = 0.69).

*Findings – Estimating Mean Achievement and Ethnic Gaps*

*Characteristics of students and schools.* The schools in our sample reflected the diversity of California’s public schools. About 9% of students enrolled were of African-American descent; 4% were Asian; 55% were Latino; and 29% were non-Latino white. Seventy-five percent of students qualify for free or reduced lunch and 34% were classified as English learners.

The mean API score for our schools equaled 676, considerably short of the state’s performance goal of 800. As shown in Fig. 1, the mean API score for the white subgroup was 734, with minimum score of 554 and a maximum of 874. Fig. 2 reports corresponding API scores for the Latino subgroup: a mean score of 663, ranging from 546 to 817.

*Multivariate results – self-reported teachers’ accountability practices.* We first examined the extent to which teacher factors accounted for variation in schools’ overall API score. Reduced models included only student

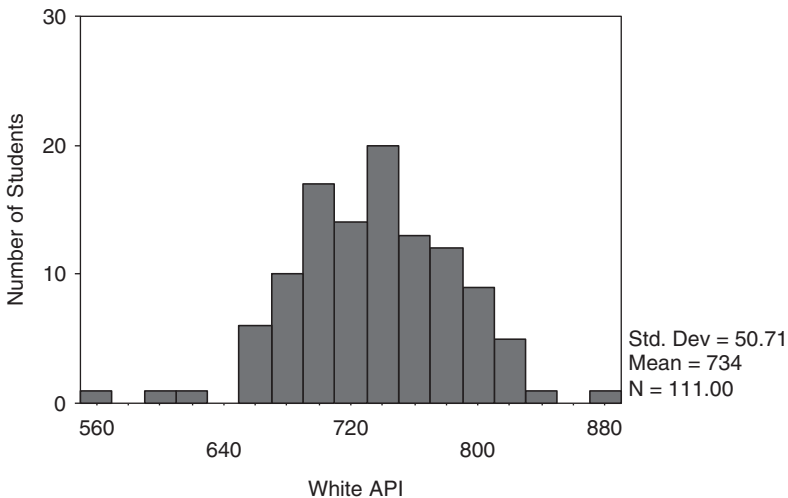


Fig. 1. Distribution of White Subgroup API Scores.



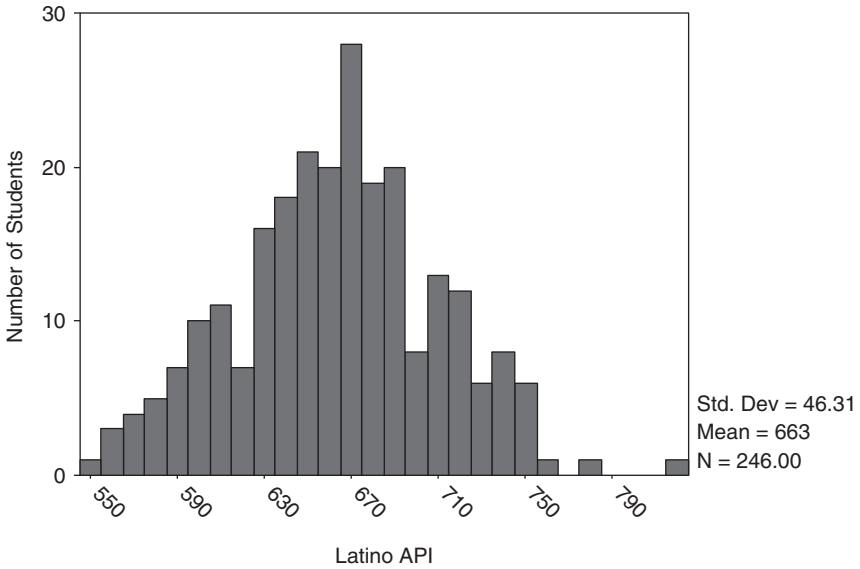


Fig. 2. Distribution of Latino Subgroup API Scores.

background (SES) variables, whereas full models included the practices reported by teachers, along with student background variables. Student background variables, as a block of predictors, explained 12% of the variance in API scores among schools.

Several student socioeconomic status (SES) factors were significant after teacher's self-reported practices were entered in the full model, including the percentage of enrollment identified as English learners, from migrant families, or geographically mobile. Teacher factors, explained another 7% of the variance (adjusted  $R^2$ ), but no individual teacher factor was statistically significant (Table 1).

Similar results emerged when estimating API scores for white and then Latino students with important exceptions. Importantly, student SES predictors explained 30% of the variance in white API scores among schools. Teacher factors account for another 7% of between school variance in white students' API scores in the full model, but no single practice as reported by teachers was statistically significant.

Mean Latino API scores among schools were less sensitive to student SES variability and a bit more sensitive to teacher-reported practices. In the reduced model, student SES predictors accounted for 12% of the variance in

**Table 1.** Estimating Academic Performance Index (API) Scores From Student Socioeconomic Status (SES) and Teacher-Reported Accountability Factors.

	School API (SE)	White API (SE)	Latino API (SE)	GAP API (SE)
<i>Reduced model</i>				
Student SES factors only				
White (%)	0.29 (0.23)	-0.64 (0.39)	-0.56 (0.26)	-0.05 (0.40)
Asian (%)	-0.03 (0.36)	-1.44 (0.75)	-0.20 (0.38)	-1.53 (0.73)
Free/reduced lunch (%)	-0.37 (0.27)	-1.93 (0.48)	-0.27 (0.29)	-1.41 (0.45)
English learners (%)	-0.60 (0.21)	-0.15 (0.37)	-0.94 (0.23)	1.25 (0.35)
Migrant (%)	-0.81 (0.31)	-0.73 (0.49)	-0.48 (0.32)	-0.12 (0.45)
Mobile (%)	-2.41 (0.70)	-1.28 (1.22)	-0.70 (0.74)	-1.23 (1.16)
Parents with high school diploma (%)	-0.23 (0.30)	-1.10 (0.47)	0.27 (0.31)	-0.78 (0.43)
Parents with college degree (%)	0.47 (0.44)	0.16 (0.78)	0.80 (0.47)	-0.59 (0.74)
R <sup>2</sup>	0.15	0.35	0.15	0.37
Adjusted R <sup>2</sup>	0.12	0.30	0.12	0.32
<i>Full model</i>				
Student SES factors				
White (%)	0.13 (0.23)	-0.28 (0.41)	-0.58 (0.27)**	-0.05 (0.45)
Asian (%)	0.01 (0.35)	-1.69 (0.72)**	-0.20 (0.38)	-1.54 (0.77)**
Free/reduced lunch (%)	-0.41 (0.27)	-1.88 (0.48)**	-0.28 (0.28)	-1.36 (0.48)**
English learners (%)	-0.44 (0.22)**	-0.01 (0.41)	-0.83 (0.24)**	0.9 (0.41)
Migrant (%)	-0.78 (0.31)**	-1.05 (0.5)**	-0.32 (0.32)	-0.31 (0.49)
Mobile (%)	-2.58 (0.68)**	-1.58 (1.28)	-0.80 (0.73)	-1.44 (1.29)
Parents with high school diploma (%)	-0.32 (0.30)	-1.18 (0.48)**	0.19 (0.32)	-0.88 (0.49)*
Parents with college degree (%)	0.51 (0.43)	0.04 (0.81)	0.92 (0.46)**	-1.04 (0.84)
Latino school 1 with high API scores	32.29 (40.43)	63.92 (42.53)	25.28 (42.50)	27.17 (41.74)
Latino school 2 with high API scores	-40.44 (40.76)	-3.63 (43.39)	-83.92 (42.75)	54.51 (42.6)

Table 1. (Continued)

	School API (SE)	White API (SE)	Latino API (SE)	GAP API (SE)
Teachers use of curriculum standards	8.6 (7.68)	2.39 (13.02)	2.03 (8.28)	7.55 (13.30)
School focuses on achievement	-0.89 (3.46)	0.97 (5.36)	-3.7 (3.74)	6.05 (5.52)
School prioritizes accountability targets	0.26 (4.10)	-6.5 (6.66)	-1.38 (4.46)	-11.57 (6.78)*
Teachers express high expectations	2.02 (5.42)	0.31 (8.92)	8.56 (6.04)	-0.07 (9.57)
District leaders prioritizes student learning	-3.04 (2.63)	4.94 (4.48)	-2.89 (2.83)	4.52 (4.76)
Teachers modify pedagogical practices	3.77 (2.46)	4.57 (3.85)	3.24 (2.61)	-2.46 (3.84)
School staff focuses on student subgroups	1.92 (3.17)	6.02 (5.15)	1.82 (3.35)	0.75 (5.11)
District staff supports teachers	2.35 (1.90)	0.71 (3.49)	0.5 (2.03)	-3.88 (3.62)
School displays a positive environment	3.59 (4.55)	4.83 (7.38)	10.74 (4.88)**	-2.66 (7.43)
Teacher development for serving ELs	-1.83 (2.28)	-0.2 (3.90)	-1.96 (2.44)	0.74 (4.05)
Constant	574.54	773.67	507.96 (81.42)	275.67 (134.01)
$R^2$	0.25	0.49	0.26	0.45
Adjusted $R^2$	0.19	0.37	0.19	0.32
$N$	245	109	237	104

\*  $p < .05$ .\*\*\*  $p < .01$ .

Latino API scores. Mean Latino API scores among schools were positively and strongly correlated with a school displaying a positive school environment. For every unit increase in this factor score, Latino API scores were almost 11 points higher, controlling for other variables. The full model accounted for 19% of the variance in Latino scores, of which 7% is attributable to teacher-reported practices.

The fourth regression model in Table 1 might be related to a school's achievement gap. The average API gap for our sample of schools was 276 points and student SES predictors accounted for 32% of the variance in mean gaps among schools. With teacher factors in the model, several socioeconomic factors were significantly and positively associated with smaller gaps in API scores ( $p < .05$ ): percent of Asian students, percent of students who qualify for free or reduced lunch, and percent of students whose parents have a high school diploma. The factor that measures whether a school prioritizes meeting accountability targets was significantly and positively associated with a smaller white–Latino achievement gap. For every unit increase in this factor score, a school's achievement gap decreased by almost 12 points, controlling for other variables. Note that negative coefficients indicate smaller API gaps. Still, teacher-reported practices did not explain additional shares of variance in white–Latino gaps.

*Multivariate results – principal-reported practices.* Next, we turn to multiple regression results for district or school practices reported by the principal. Overall, API scores and gaps were more sensitive to principal reports of effective practices, compared with teacher reports, although student SES continued to be more strongly related to achievement. Two principal factors were significantly and positively associated with mean API scores among schools. For every one-unit increase in the factor score that gauges the strength of districts' guiding a school's curriculum, as reported by the principals, mean API scores were almost 3 points higher, controlling for other variables. For every one-unit increase in the factor score tapping whether teachers express high expectations, a school's API score increased by 3 points, controlling for other variables. In this model, student socioeconomic variables and principal factors explained 21% of the variance in a school's API score. The inclusion of principal factors led to an additional reduction of variation of 10% in a school's API score (Table 2).

Next, we report results for estimation of mean API scores for white and then Latino students among schools. Our final principal models examined factors that might be related to student achievement for our two distinct subgroups. The average API score for the white subgroup was 758 and

**Table 2.** Estimating Academic Performance Index (API) Scores From Student Socioeconomic Status (SES) and Principal-Reported Accountability Factors.

	School API (SE)	White API (SE)	Latino API (SE)	Gap API (SE)
<i>Reduced model</i>				
Student SES factors only				
White (%)	0.29 (0.23)	-0.64 (0.39)	-0.56 (0.26)	-0.05 (0.40)
Asian (%)	-0.05 (0.36)	-1.44 (0.75)	-0.20 (0.38)	-1.53 (0.73)
Free/reduced lunch (%)	-0.66 (0.27)	-1.93 (0.48)	-0.27 (0.29)	-1.41 (0.45)
English learners (%)	-0.35 (0.21)	-0.15 (0.37)	-0.94 (0.23)	1.25 (0.35)
Migrant (%)	-0.85 (0.31)	-0.73 (0.49)	-0.48 (0.32)	-0.12 (0.45)
Mobile (%)	-1.35 (0.71)	-1.28 (1.22)	-0.70 (0.74)	-1.23 (1.16)
Parents with high school diploma (%)	0.04 (0.30)	-1.10 (0.47)	0.27 (0.31)	-0.78 (0.43)
Parents with college degree (%)	0.46 (0.45)	0.16 (0.78)	0.80 (0.47)	-0.59 (0.74)
$R^2$	0.14	0.35	0.15	0.37
Adjusted $R^2$	0.11	0.30	0.12	0.32
<i>Full model</i>				
Student SES factors				
White (%)	0.17 (0.23)	-0.30 (0.41)	-0.42 (0.27)	-0.18 (0.40)
Asian (%)	0.07 (0.38)	-1.68 (0.79)**	-0.25 (0.41)	-1.73 (0.72)**
Free/reduced lunch (%)	-0.35 (0.28)	-1.57 (0.50)**	-0.24 (0.31)	-0.63 (0.49)
English learners (%)	-0.32 (0.22)	0.27 (0.38)	-0.66 (0.25)**	1.155 (0.36)**
Migrant (%)	-0.57 (0.31)*	-0.68 (0.50)	-0.25 (0.33)	-0.69 (0.46)
Mobile (%)	-2.87 (0.71)**	-3.06 (1.27)**	-1.30 (0.77)*	-2.25 (1.18)*
Parents with high school diploma (%)	0.13 (0.30)	-0.58 (0.49)	0.55 (0.32)*	-0.86 (0.45)*
Parents with college degree (%)	0.37 (0.46)	-0.15 (0.82)	0.61 (0.50)	-0.20 (0.79)
Latino school 1 with high API scores	44.47 (40.66)	70.92 (43.20)	40.21 (43.15)	27.91 (39.30)

Principal-report accountability factors	Latino school 2 with high API scores	-51.88 (40.50)	-29.93 (42.56)	-98.08 (43.03)	61.64 (38.75)
	School focuses on achievement	2.70 (2.33)	-0.63 (3.34)	0.91 (2.48)	1.90 (3.04)
	District guides school curriculum	2.72 (1.47)*	1.10 (2.31)*	4.31 (1.58)*	0.26 (2.14)
	Teachers express high expectations	3.01 (1.78)*	1.89 (2.63)**	3.55 (1.93)**	0.06 (2.46)
	District expects schools to meet account.	3.47 (3.07)	2.24 (4.88)	2.42 (3.40)	-5.52 (4.61)
	School prioritizes accountability targets	0.61 (2.48)	0.07 (3.63)	1.50 (2.66)	-2.66 (3.32)
	District expects schools to improve	-0.20 (2.72)	6.71 (4.95)	-2.01 (3.00)	1.04 (4.73)
	School staff focuses on student subgroups	-0.32 (1.16)	-0.10 (1.91)	-1.17 (1.24)	-1.86 (1.76)
	School displays a positive environment	1.87 (2.02)	3.35 (3.01)	0.85 (2.17)	0.75 (2.81)
	Constant	564.25 (51.49)	758.33 (83.26)	537.32 (57.17)	258.90 (79.90)
	$R^2$	0.27	0.51	0.26	0.51
	Adjusted $R^2$	0.21	0.39	0.19	0.39
	N	217	97	209	93

\*  $p < .05$ .

\*\*  $p < .01$ .

student socioeconomic variables explained 30% of the variance in white API scores. When principal factors were added to the model, several socioeconomic factors were significantly and negatively correlated with white API scores ( $p < .05$ ): percent of Asian students, percent of students who qualify for free or reduced lunch, and percent of mobile students.

The same principal factors associated with a school's API score, district guides a school's curriculum and teachers express high expectations, were also associated with the white API score. For every unit increase in district guiding a school's curriculum, the white API score increased by 1 point. For every unit increase in teacher expectations, the white API score increased by almost 2 points, controlling for other variables. In the full model, student socioeconomic variables and principal factors explained 39% of the variation of the white API score. The inclusion of principal factors led to an additional reduction of variation of 9% in the white API score.

When estimating mean API scores for Latino students, student SES accounted for 12% of the variance. As was true with the model for school's overall API score and the white API, districts guide a school's curriculum and teachers express high expectations, were also associated with the Latino API score. For every unit increase in district guidance, the Latino API score increased by 4 points, controlling for other variables. For every unit increase in teacher expectations, the Latino API score increased by almost 4 points, controlling for other variables. In this model, student socioeconomic variables and principal factors explained 19% of the variation of the Latino API score. The inclusion of principal factors led to an additional reduction of variation of 7% in the Latino API score.

Our final model examined principal factors that might be related to a school's achievement gap. The average achievement gap for our sample of schools was 259 points, and socioeconomic variables explained 32% of the variance in a school's API score. When principal factors were added to the model, several socioeconomic factors were significantly and positively associated with closing a school's achievement gap ( $p < .05$ ): percent of Asian students, percent of mobile students, and percent of students whose parents have a high school diploma. The percent of EL students was significantly and negatively associated with increasing a school's achievement gap. None of the principal factors were significantly correlated with closing a school's achievement gap. In this model, student socioeconomic variables and principal factors explained 39% of the

variance in a school's achievement gap. The inclusion of principal factors led to an additional reduction of variation of 7% in a school's achievement gap.

## **DISCUSSION AND POLICY IMPLICATIONS**

While this study analyses a sample of schools that were selected to minimize the correlation between student socioeconomic characteristics and student achievement, our regressions indicate that these variables still play a role in determining school achievement. Reviewing the teacher regressions, no teacher factor plays an important role in improving a school's API scores; meanwhile, two student socioeconomic variables are significant. The same trend holds for teacher factors and a school's achievement gap, as only one teacher factor is significantly correlated to closing the achievement gap; three student socioeconomic variables are significant. Regarding the principal regressions, two principal-level factors are correlated with a school's API score; two student socioeconomic factors are significant. Furthermore, no principal factors made a significant impact on a school's achievement gap; four student socioeconomic variables are significantly correlated with this outcome measure.

We have seen how teachers, principals, and districts play a role in shaping achievement. However, the broader implication is that a school's overall API test score, achievement gap, and subgroup scores remain linked with student socioeconomic variables. California's current accountability model is built upon the foundation that through the dedicated use of standards and assessments and clear and definitive consequences, schools will have the capacity and ability to improve achievement. Our findings imply that API scores and achievement gaps tend to be susceptible to student level characteristics, variables that schools and districts have no control over. The relationship between student socioeconomic factors and school achievement is a reminder that current accountable models may fall short in helping schools transform and improve student achievement.

Another implication of our study is that current accountability tools, as measured by our factors, might have a slightly larger impact on Latino students. It appears that accountability is more effective, in a sense, for Latino students. The reasons for this could be twofold. First, the white students in our sample already have relatively high API scores, with an average API score of 734. With a designated API score of 800, the white



students in our sample appear on track to reach this goal. However, the API score for Latino students in our sample is much lower, as their average API score is 663, a difference of over 70 points from their white counterparts. In this regard, white students do not have as much distance to improve, and current accountability policies might not be associated with raising their level of achievement.

Another plausible explanation is that accountability, which places an increased focus on improving achievement among all students, has pushed schools in our sample to increase their efforts on reaching traditionally underserved or underperforming students. While this set of accountability policies and subsequent efforts have not yet substantially helped schools close the achievement gap, several factors did help schools raise the API test scores for their Latino students. In this regard, Latino students appear to benefit from current first-wave accountability tools and policies. While it appears promising that Latino students might be more likely to benefit under standards-based accountability, there is one caveat. Our teacher and principal regression models explain 37% and 39% of the variation in white API subgroup scores. However, the same regression models only explain 19% of the variation in Latino API subgroup scores, a considerably smaller amount. Perhaps our models, which are fairly strong predictors for the achievement of white students, are inherently weaker predictors for Latino students. There could be other factors or variables not included in this study that might have a stronger impact on Latino API scores.

Finally, our findings suggest that differing policy aims, raising a school's overall student performance versus narrowing the achievement gap, might require different policies. California's current accountability policies compel schools to focus on raising student achievement across the board. Under this framework, increasing achievement across all subgroups is the desired outcome. Equity is interpreted as the upward trend of all students. Can a school use the same tools to close their achievement gap? As revealed in our regressions, a school's API score and achievement gap appear to be correlated with different factors. As shown, a school's API score is positively impacted by two factors: teacher expectations and district guidance regarding curriculum. In this sense, API scores in our sample are correlated with two straightforward findings: schools in which teachers communicate and hold students to high expectations, and are provided with a clear and coherent and aligned curriculum, tend to have higher API scores. It appears that a school's API score is linked to school (teacher expectations) and district (support of curriculum) factors that might

ultimately influence the black box of classroom instruction, perhaps leading to noticeable improvement on a school's API score.

While our regression findings indicate that a school's API score is associated with various accountability tools experienced by local educators, it is disappointing to find that none of these factors correlate with helping schools close their achievement gap. In fact, a school's achievement gap strongly correlates with only one factor: teachers report that their schools made a conscious decision to prioritize and focus on both school and subgroup accountability targets. The incongruity of our findings lends weight to an overall argument that while certain accountability mandates have been helpful in improving a school's overall API score, these tools might not be as effective in helping schools close the achievement gap between white and Latino students. Furthermore, the paucity of findings for the achievement gap regressions also lay cause to the notion that factors that would help schools close the achievement gap are not included in our regression analyses.

The original research project was designed to identify teacher, school, and district factors that helped schools "beat the odds," thereby enabling them to have higher than expected API scores. This singular focus could play a role in limiting the creation of factors that might have been more related to closing the achievement gap. However, survey questions did touch upon numerous policies and practices that, in theory, should play a role in helping schools close achievement gaps. Overall, our findings indicate that closing the achievement gap might require additional policies and practices that are specifically targeted to increasing the achievement of selected subgroups.

Overall, this study demonstrates that scholars working in the effective schools traditions must recognize that accountability tools and practices can affect teachers' and principals' everyday work. The fact that these accountability policies are observed by local educators and, at times, are related to achievement levels confirm Rowan's (2006) recent argument that the loose-coupling metaphor must be rethought in the context of top-down accountability. The penetration of accountability practices into the classrooms technical core now challenge the earlier imagery of loose coordination of work across organizational levels. Yet the effects of reported practices – whether seen through the eyes of teachers or principals – were modest in magnitude. The tidy transmission of learning standards, aligned pedagogy, and performance monitoring intended by policy makers does not appear to uniformly impact organizational actors. Rather, it appears that while teachers and principals feel these new demands and resources, they have an uneven effects on their everyday work, consistent with the findings of other chapters in this volume.

As revealed in our analysis, schools play a significant role in improving student achievement. It is not clear, however, that current accountability policies, as measured in our model, help schools tackle the intricate and complex problems that might contribute to the achievement gap. Nor do these policies eclipse the prior force of students' backgrounds, even after truncating variation in family social-class through our innovative design.

Another conclusion is that current accountability tools, as measured in our model, have a narrow purpose, as they appear to help schools improve their overall student achievement. Yet these policies exhibit limited efficacy in closing achievement gaps. Our study suggests that if schools are to move towards a new definition of equity, one that requires achievement parity across all subgroups, schools must have access to a different set of policies and practices. Until then, the achievement gap remains a powerful reminder of the difficulty in achieving and maintaining equity in our public schools.

## NOTES

1. The original study and further design details can be found at [http://www.edsource.org/pub\\_abs\\_simstu05.cfm](http://www.edsource.org/pub_abs_simstu05.cfm)
2. These constant-school-sample regression results are available from the authors.
3. In the overall sample of schools, only 35 schools had API scores for an Asian subgroup, and 54 schools had African-American subgroup API scores.

## ACKNOWLEDGMENTS

The authors would like to thank Betsey Woody, Anne Driscoll, Leah Walker, and Willow Sussex for their various contributions in helping us analyze, frame, and understand our complex data and findings. The *Similar Students, Differing Results* study was led by Ed Haertel, Mike Kirst, and Trish Williams. Bruce Fuller was helpful in providing guidance throughout this project, and we thank him for generous feedback and assistance. Any errors or omissions are those of the authors.

## REFERENCES

- Bell, J. (2001). High performing, high poverty schools are distinguished by their strength of leadership and their understanding of how children learn. *Leadership*, 31(1), 8–11.

- Carter, S. C. (2000). *No excuses: Lessons from 21 high-performing, high-poverty schools*. Washington, DC: The Heritage Foundation.
- Cawelti, G. (1999). *Portraits of six benchmark schools: Diverse approaches to improving student performance*. Arlington, VA: Educational Research Service.
- Education Trust. (1999). *Dispelling the myth: High-poverty schools exceeding expectations*. Washington, DC: Council of Chief State School Officers/U.S. Department of Education.
- Fuller, B., with Bridges, M., & Pai, S. (2007). *Standardized childhood: The political and cultural struggle over early education*. Palo Alto: Stanford University Press.
- Izumi, L., with Coburn, K., & Cox, M. (2002). *They have overcome: High-poverty, high-performing schools in California*. San Francisco: Pacific Research Institute.
- Johnson, J. (1999). *Hope for urban education*. Austin, TX: The Charles A. Dana Center, University of Texas at Austin.
- Levine, D. U., & Lezotte, L. W. (1990). *Unusually effective schools: A review and analysis of research and practice*. Madison, WI: The National Center for Effective Schools Research and Development.
- Meyer, J. W., & Rowan, B. (1977). Institutional organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83, 340–363.
- Meyer, H. D., & Rowan, B. (2006). Institutional analysis and the study of education. In: H. D. Meyer & B. Rowan (Eds), *The new institutionalism in education* (pp. 1–14). Albany: State University of New York Press.
- Perry, M., & Pelavin, A. (2003). *California's lowest-performing schools: Who they are, the challenges they face, and how they're improving*. Palo Alto, CA: EdSource.
- Phi Delta Kappan (2001). A Kappan special section on school reform. *Phi Delta Kappan*, 83(1), 59–95.
- Public Schools of North Carolina. (2000). *Closing the achievement gap: Views from nine schools*. Raleigh, NC: State Board of Education/Department of Public Instruction.
- Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *The Elementary School Journal*, 83(4), 426–452.
- Rosenholtz, S. J. (1985). Effective schools: Interpreting the evidence. *American Journal of Education*, 93(3), 352–388.
- Rowan, B. (2006). The new institutionalism and the study of educational organizations: Changing ideas for changing times. In: H. D. Meyer & B. Rowan (Eds), *The new institutionalism in education* (pp. 15–32). Albany: State University of New York Press.
- Smith, M. S., & O'Day, J. A. (1991). Systematic school reform. In: S. Fuhrman & B. Malen (Eds), *The politics of curriculum and testing* (pp. 233–267). New York: The Falmer Press.
- Virginia Department of Education. (2000). *A study of effective practices in Virginia schools: Educators' perspectives of practices leading to student success*. Richmond, VA: Division of Accountability, Governor's Best Practice Centers.
- Williams, T., Kirst, M., & Haertel, E. (2005). *Similar students, different results: Why do some schools do better? A large-scale survey of California elementary schools serving low-income students*. Mountain View, CA: EdSource.

## CHAPTER 6

# HIGH STAKES DIPLOMAS: ORGANIZATIONAL RESPONSES TO CALIFORNIA'S HIGH SCHOOL EXIT EXAM

Jennifer Jellison Holme

### ABSTRACT

*This chapter examines organizational and instructional responses of California's high schools to the introduction of a High School Exit Examination through interviews with 47 high school principals across the state. I found that most schools changed little about their organizational structure, and provided little support for students until after they failed the exam. Findings also indicate that the exit exam influenced the curriculum most significantly in low-performing schools and in low-track classes within higher performing schools. While the exit exam spurred some positive changes, it also led to unintended consequences inside classrooms.*

### INTRODUCTION

State policymakers are increasingly turning to high-stakes exit examination requirements to hold high schools and their students to higher academic

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
Research in Sociology of Education, Volume 16, 157-188  
Copyright © 2008 by Emerald Group Publishing Limited  
All rights of reproduction in any form reserved  
ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16006-2

standards and to “restore value” to high school diplomas (Achieve, 2005). As of 2007, 22 states required students to pass an “exit exam” in English/Language Arts and Mathematics (and in some states, tests in additional subject areas) to obtain a diploma. Many of these states are currently in the process of upgrading their exam systems by increasing the number of subject areas tested and raising their exam’s skill levels. All together, these 22 states enrolled 65% of the nation’s public school students, and 76% of the nation’s minority students (CEP, 2007, p. 1). Should the four additional states that have planned to institute exit examination requirements go forward, by 2012 exit exams will affect 76% of all public school students, and 82% of all minority students in the U.S. (CEP, 2007, p. 1).

Exit exam policies embody a number of assumptions about the kinds of incentives and pressures that best motivate educators and students. In theory, the high stakes tests are intended to prompt significant changes in the way schools behave organizationally by forcing schools’ attention onto their lowest-performing students. At the same time, the sanctions attached to the exams are hoped to motivate both educators and students to increase their work efforts (Darling-Hammond, Rustique-Forrester, & Pecheone, 2005).

While exit examinations have been in effect in some states for over two decades, little is known about how schools respond to these mandates. To date, most of the research on exit examinations has focused on the impact of exit exams on measures of attainment and achievement (see, e.g., Amerin & Berliner, 2002; Greene & Winters, 2004). Few researchers have detailed the extent to which schools change as organizations in terms of the pedagogical supports, curriculum, and instruction they provide to students. The handful of more in-depth studies that do exist have relied on limited data, and then have aggregated data from large cross-sections of schools with little attention to local contexts (see, e.g., CEP, 2007; HumRRO, 2006).

This chapter applies a new framework to research on exit exams by examining the *variation* in schools’ responses to exit examination requirements across different schooling contexts (Diamond & Spillane, 2004). This chapter also adds an important dimension to the existing school reform literature that tends to examine organizational and instructional change in isolation from one another. The present study not only documents the range of structural supports that schools offer for students and the ways in which curriculum changes in response to an exit exam; I also consider how curriculum changes in the presence – or absence – of the structural supports that are (or are not) put in place.

## BACKGROUND AND HISTORY OF EXIT EXAMS

Exit exams have been a part of state testing policy for over three decades. In the 1970s and early 1980s, most states adopted exams – often called minimum competency tests (MCTs) – that students were required to pass in order to receive a diploma. While these tests were somewhat different than today’s more rigorous standards-based exit exams – as the MCTs were intended to ensure high school graduates had mastered only minimum basic skills, often at the 8th grade level – the MCTs do share some political and economic origins with the exams of today in that the policies emerged out of some similar economic concerns about the migration of U.S. jobs overseas.

When the MCTs were instituted in the mid to late 1970s, the U.S. economy was saddled with a “trifecta” of economic troubles: high inflation, economic recession, and high unemployment. During this time, corporations began moving manufacturing plants overseas to countries with lower labor costs, sending the secure and well-paying middle class jobs with them. Policymakers, believing that low educational standards were at least partly to blame, responded by adopting MCT requirements to ensure that high school students were graduating with minimum basic skills. Between 1973 and 1983, the number of states with MCTs rose from 2 to 34 states – 12 more than currently have exit exams today. These tests marked a turning point from the use of large-scale assessments for program evaluation to the use of assessments for school accountability (Hamilton & Koretz, 2002; Heubert & Hauser, 1999, p. 15).

This focus on minimum skills tests shifted in 1983 with the publication of “A Nation at Risk” – a report calling for higher standards and more rigorous curriculum to shore up national competitiveness at a time when the “deindustrialization of America” was understood as a permanent trend. The “new” economy, the report argued, would require that students master higher-level skills (Beyer, 1985; Hamilton & Koretz, 2002). Policymakers responded by instituting tougher graduation requirements and increasing the rigor of high school classes (Tyack, 1993).

This most recent wave of exit exam policies – enacted by laws passed in the late 1990s and early 2000s – can be traced (as with MCTs) to economic anxieties, although these new policies were not driven by recessionary concerns. Many state exit exam policies were, in fact, enacted at a time of unprecedented economic growth, prior to the collapse of the dot-com bubble in 2000. Yet underneath the economic upturn was forebidding news about the increased off-shoring of professional, white-collar positions to China, India and Singapore, a trend begun in the late 1990s. Policymakers’ response, as in the 1970s, has been to seek an “education solution” to uneven

economic trends, and the result has been the re-adoption of high school exit examination policies. These policies differ from their MCT predecessors in that they focus (or claim to focus) on higher-order thinking skills, and are set to higher standards (at least 10th grade standards). Much of the rationale for these policies has been centered on the need for better-educated workers in a global economy. As California's State Superintendent Jack O'Connell claimed, in defending his exit exam policy: "We face a new economy driven by global innovation that will demand higher-level skills and knowledge to meaningfully enter the work force. It is imperative that all of California's children reach at least the minimal bar set by our exit exam" (CDE, 2006).

There is, however, growing evidence that current global economic trends affecting U.S. job growth have deeper roots than the educational system. In fact, some economists argue that it is not low skills that are to blame for current economic trends, but a lack of jobs fostered by national policies that do little to prevent the off-shoring of higher-paying jobs (Mishel, Bernstein, & Allegretto, 2006). These trends, economists observe, have transformed the U.S. economy from an "hourglass shape" of the late 1990s (with few middle class jobs but adequate jobs at the upper- and lower-middle levels) to a "barbell" shape in which the greatest expansion in jobs (and income) is occurring only at the two extremes. As a Trumbull (2006) writes, "during the current expansion [in early 2006], the bulk of new jobs have come in either the highest-paid of five broad occupational categories – management and professional – or the lowest-paid, services. Together the two sectors now account for more than half of all jobs" (p. 1). This decline in upper-middle-income paying positions has led to declining real wages and increasing unemployment for college graduates in recent years (Gosselin, 2006).

Students who lack high school diplomas, therefore, face a declining future, as the likelihood of obtaining a well-paying middle class job is less likely today than it was 30 years ago at the height of the MCT movement. Indeed, the "achievement gaps" between racial, ethnic, and socioeconomic groups have become much more consequential in terms of earnings in recent years. Despite declining wages and employment for college graduates, college graduates still earn on average 80% more than high school graduates, while high school graduates are four times more likely to be unemployed (Olson, 2006).

Policy makers enacted high school exit examination policies in response to these trends, as a mechanism to shore up the skills of the lowest achieving high school students. The important policy question becomes, first, high school exit exams like California's bring about hoped changes in the behavior of schools and students? But even more significantly, and what this chapter explores, is under what conditions, and to what effect?



## UNDERLYING POLICY ASSUMPTIONS OF EXIT EXAMS

Though rarely if ever specified, in theory, exit exams are supposed to increase student achievement through multiple “pressure points” (Mintrop, 2003; O’Day, 2002). First, exams are supposed to *increase the “motivation” of students* by creating a concrete hurdle that they must clear prior to high school graduation (Dee & Jacob, 2006). Under this assumption, the problem of low achievement is conceptualized as one of lack of effort or will on the part of the students – and thus a high stakes exam is seen as a way to force students to increase their effort and take responsibility for their learning. (Theoretically parents could also be expected to pay more attention to the educational process by this requirement as well.)

Exit exams are also intended to prompt *change in organizations*, first through “naming and shaming” (Mintrop, 2003) as school-wide scores on exit exams are made public; second, through threats of sanctions: 20 of the 22 states with current exit exams tie school scores to No Child Left Behind (NCLB) accountability (CEP, 2006, p. 80). Exit exams are, therefore, intended to prompt schools as organizations – individual educators as a *collective* – to undertake systemic structural and instructional change (Elmore, 2003). Though rarely specified, school change is also assumed to be a matter of degree: schools with large numbers of failing students would be expected to undertake more significant organizational change, whereas schools with fewer lower achieving students would be expected to target their efforts on the lowest achieving students in particular.

To accomplish these organizational changes, exit exams are intended to prompt changes in the *individual behavior of actors within the organization*, including teachers, principals, and counselors (Mintrop, 2003). One means of creating the motivation for change is through threat of sanction; another is through appeals to professional pride and desire to improve (Mintrop, 2003).

An important mechanism by which these changes are supposed to be accomplished is through educators’ *increased utilization of the information* that these exams provide. This information is intended to drive instructional improvement both by identifying low-achieving students, and by highlighting content areas that need attention (Muller & Schiller, 2000; Rhoten, Carnoy, Chabran, & Elmore, 2003).

Exit examination policies, therefore, implicitly rely on the motivational power of sanctions together with test results to accomplish large-scale educational change. Capacity building (resources, professional

development, etc.) is not a main focus of these laws. Further, these policies assume schools that will look to themselves or their districts to figure out “what to do” (Elmore, 2003). The knowledge is assumed to exist, and it is up to educators to find and implement the correct strategies for their particular schools.

It is important to note that the “pressure” on high schools that flow from exit examination policies is distinct – and greater – than NCLB accountability pressures alone, in terms of *who* is being held accountable, *when* this accountability occurs, and how accountability is measured. The first and most significant pressure point comes at the first time the schools administer the exam, because it is this test administration that is used (in most states) for NCLB accountability; thus, scores for this “census” administration matter a great deal. Additional pressure is placed on schools to provide assistance and support for students who fail the first administration of the test, yet this pressure is largely informal pressure in that it comes from students, parents, and the public pressure of published cumulative pass rates in those areas in which newspapers publish 12th grade results. Exit examination policies, therefore, put accountability pressure on multiple levels of the educational system, on multiple actors within that system, at multiple junctures of a student’s career.

Understanding how schools respond to the California High School Exit Examination (CAHSEE) requirement is not a study of policy implementation, as schools were given no specific policy to implement. The test *is* the policy: school change is expected to flow backwards from the test. How that change happens and what types of changes are made are entirely up to the discretion of educators within the school and district. As such, this is not a study of policy implementation but rather a study of *policy response* to a high-stakes test (see also Elmore, 2003).

## EXISTING RESEARCH

While research on high schools’ response to high-stakes testing has increased in recent years (see, e.g., Carnoy, Elmore, & Siskin, 2003; Diamond & Spillane, 2004; Mintrop, 2003), very few researchers have studied schools’ organizational and instructional responses to exit exam policies specifically. Existing studies have yielded quite mixed results, in large part because they fail to attend to the ways in which exit exams play out differently across different types of schooling contexts. Some of these studies – particularly studies of “beat the odds” schools – have found that

high stakes tests can encourage a focus on state standards and lead to increased content coverage across classrooms, can prompt schools to provide increased supports for low achieving students, and that such policies can often prompt the reassignment of better teachers to lower achieving students (Gayler, 2005; HumRRO, 2006; Skra, Scheruich, Johnson, & Koschorek, 2004). Other studies – especially those studies that have aggregated data from a variety of schools – have found that not all responses to exit examination policies are positive: some studies have documented curricular narrowing for disadvantaged students, an increased focus on test-taking skills, and a decreased emphasis on untested subjects (Gayler, 2005; McNeil & Valenzuela, 2001).

This chapter brings an important new perspective to the exit exam literature by examining the variation in schools' responses to exit examination requirements across different schooling contexts (Diamond & Spillane, 2004). This chapter also adds an important dimension to the existing school reform literature that tends to examine organizational and instructional changes in isolation from one another, by analyzing where organizational and instructional changes intersect. This chapter therefore not only documents the range of structural supports schools offer for students, or the ways in which curriculum changes in response to exit exam mandates, it also considers how curriculum changes in the presence – or absence – of the structural supports that are (or are not) put in place.

## **METHODS AND RESEARCH QUESTIONS**

This study was guided by the following research question: What kinds of organizational and instructional responses do California's high schools make in response to the state's exit examination requirement? In 1999, California passed legislation requiring that all students pass the CAHSEE in order to obtain a diploma, beginning with the Class of 2004. The 10th grade scores on the exam, which consists of two components – an English Language Arts (testing standards through 10th grade) and a Mathematics portion (testing students through Algebra 1) – are used for NCLB accountability for high schools.

A total of 47 high schools were selected for this study using stratified random sampling based on the Spring 2004 test administration, which was the first administration of the CAHSEE for the Class of 2006, and the first time the current version of the test was used. Schools were sampled from six

California counties, selected for diversity in urbanicity, size, and location. These counties, taken together, included 53.2% of the state's student population. Schools were sampled within counties, weighted by county population; thus, more schools were sampled in the larger counties to make the final sample of schools somewhat approximate to that county's portion of the state population.

Schools were randomly sampled within these counties for their initial test score performance to capture as broad a range of schools as possible. Sampling categories were based on a previous analysis of state data conducted by Rogers, Holme, and Silver (2005), which analyzed the performance of California's high schools and divided the schools into quartiles by performance. Based on this analysis, the schools were folded into three categories: low pass rate (LPR) schools, with pass rates below 70% on *either* the math or ELA portion of the exam in spring of 2004, housing roughly 25% of the state's students; moderate pass rate (MPR) schools, with pass rates between at least 70% and 90% on one or both portions of the test, housing roughly 50% of the state's students; high pass rate (HPR) schools, with more than 90% initial pass rates on *both* portions of the state exam, housing the last 25% of the state's students. Ultimately, the sample included 47 high schools, 25 of which (53%) were LPR schools, with pass rates of below 70% on either the math or ELA portion of the exam; 14 (or 30%) of the schools in the sample were MPR schools with between 70% and 90% passing both portions of the test; and 8 (or 17%) of which were HPR schools with more than 90% initial pass rates on both portions of the test. Six of the schools in the study (13%) were "stand-alone" small schools, as they were housed on separate campuses (as opposed to small learning communities within larger comprehensive high schools); two of these were charter schools. All of the small schools/charter schools were considered low performing and were LPR schools.

Interviews were conducted with high school principals, lasting between 40 and 90 minutes each. Eleven of the interviews were conducted over the telephone. The interview protocol included questions about the principals' background and tenure in the district; curriculum and instructional change; organizational and programmatic change; professional development; state and district-level support; test administration and logistics issues; identification and notification of students at-risk of not passing; resource questions; supplemental supports for students; and dropout and retention issues. Interviews were fully transcribed and coded for cross-case themes.

The findings described here are limited to self-reports by school leaders, who are in a unique position to reflect about the way in which the CAHSEE

has impacted their school on a variety of dimensions. While principal data is limited in that the instructional and curricular changes they describe cannot be triangulated, it is likely that principals under-reported negative impacts or consequences of the test on organizational supports and curricular and instructional change. Furthermore, understanding how principals – who are held personally accountable for schools’ improvement – make sense of their schools’ change process sheds important insight on the influence of accountability pressures that stem from high-stakes testing policies. These reports are retrospective and interpretive; thus, they consist of principals’ understandings of both the character and timing of their schools’ response to the CAHSEE. In the text of this chapter, as shall be noted below, in some instances school change had occurred prior to the CAHSEE; care is taken to note such instances in the results and analysis.

It is important to note that the outcomes in this study are not defined as test scores; as such this study is not examining which schools were more successful at getting students to pass the CAHSEE itself. This study is also not aimed at describing the “average” response of all California’s high schools to the CAHSEE requirement by generalizing these findings to all high schools in the state. Rather, through the use of in-depth interviews, this study documents the way in which schools across different contexts responded both organizationally and in terms of curriculum and instruction to this requirement, to better understand the impact of these changes on the education provided to the lowest achieving students within particular schooling contexts. In this way, this study is one of the first to look beyond a set of low achieving schools, or even “beat the odds” high poverty schools, to better understand how this reform interacts with the state’s diverse array of schools; and to understand how low achieving students within these differing schooling contexts are – or are not – provided with different types of schooling opportunities as a result of this mandate.

This study pays particular attention to the ways in which schools at differing initial performance levels (in terms of the 10th grade administration of the CAHSEE in Spring of 2004, the first time the class of 2006 took the exam) respond to the exit exam requirement. According to the implicit theory of school change embedded in these high-stakes testing policies, schools *should* theoretically respond differently to exit exam requirements based on their initial overall student scores: for example, schools with large numbers of failing students would be expected to interpret their poor test scores as a mandate to undertake extensive changes in their school in terms of organization and curriculum. By contrast, schools with fewer failing students would hypothetically engage in targeted and focused support on

their lowest achieving students— students who were perhaps previously overlooked prior to the accountability policy. Yet the vast majority of school reform literature rarely pays attention to schools’ “starting points,” instead aggregating data across schools with the assumption that where schools started in terms of student performance makes little difference in their response to the policy. This study, by contrast, takes schools’ initial starting points as the foundation of analysis: school response is considered to depend critically on where they began. Thus, this chapter not only examines overall trends across all schools, it takes a close look at the differences in responses between schools with low initial performance levels and schools with higher initial performance levels on the exam.

The CAHSEE presents an unusually opportune moment to study this type of differentiated response, because the test was re-vamped and re-set for the Class of 2006, which first took the test as sophomores in the Spring of 2004. The changes for Spring of 2004 included a significant reduction in questions on the ELA section from the previous year, and the elimination of some of the advanced mathematics problems (HumRRO, 2004, p. 13). These changes were significant enough to make the scores prior to 2004 incomparable; thus, this study takes as its starting point the version of the test that has been used in all subsequent years, which was the version administered in spring of 2004 to the Class of 2006 as 10th graders.<sup>1</sup>

California’s high schools, of course, were not “tabula rasa” in spring of 2004 when the new CAHSEE was administered; educators had received, for a number of years, a great deal of feedback about school and student performance through the state’s other standardized tests (the California Standards Tests (CSTs) in grades 9 through 11, and previously, the Stanford-9; as well as older versions of the CAHSEE). The CAHSEE that was presented to schools in 2004 was, however, a new yardstick with higher stakes for students and for schools under NCLB.

This study pays attention to the 10th grade scores, in that the test is administered to all students for the first time in the spring of 10th grade. Importantly, these 10th grade census scores are used for NCLB accountability. Thus, educators would be theorized to pay close attention to these 10th grade scores and craft their responses accordingly. The cutoff scores for NCLB accountability are set at higher than the score required for passing the test for graduation.

While the primary analytic category that drove much of this analysis was the school’s initial performance on the spring 2004 CAHSEE administration for 10th graders in the Class of 2006, it is important to note that these schools vary on other dimensions: because race and poverty are closely

correlated with testing outcomes, it is not surprising that many schools that fell within the “Low Pass Rate” category were urban high schools that served large numbers of low income students of color. Yet, as the data illustrates below, school’s responses oftentimes defied what might be initial predictions based on demographics and location.

### **ORGANIZATIONAL AND INSTRUCTIONAL RESPONSES: LIMITED SUPPORT, NARROWING THE CURRICULUM**

Understanding organizational change in schools entails an examination of the ways in which schools alter – or elect not to alter – their school structures (such as school schedules, student grouping practices), and “the way content or subject matter of education is allocated time during the school day” (Elmore, Peterson, & McCarthy, 1996, p. 2). Researchers have established that organizational changes are neither necessary nor sufficient to generate change in curriculum or instruction (Elmore, Peterson, & McCarthy, 1996). As Fullan (2001) has concluded, it is often possible to change schooling structures without changing much of anything that happens in the classroom as far as the use of different materials, teaching strategies, or – often more importantly – teachers’ beliefs and assumptions about both their teaching practice and about the students that they teach (Fullan, 2001, p. 39, see also Elmore, Peterson, & McCarthy, 1996 and Oakes, Wells, Jones, & Datnow, 1997). Because the data from this study are drawn solely from principals, the focus of this analysis is on school structures and instructional changes (as self-reported). This chapter, therefore, examines organizational and instructional changes (as described by principals) independently, and concludes with an examination of the ways in which the two intersect.

#### *Organizational Change: Marginal Changes, Ad Hoc Supports*

The 47 principals interviewed described a vast array of organizational responses to the CAHSEE in terms of scheduling changes, course changes, and support structures for students.<sup>2</sup> For the purposes of analysis, their responses were categorized into three broad analytic categories, ranging from the least organizational change and supports for students (labeled

Level 1 Support), to slightly more significant change (labeled Level 2 Support) to the greatest support (Level 3 Support). The majority of schools in this study fell into Level 1 and Level 2 Support categories – engaging in relatively little organizational change in response to the CAHSEE (36 of 47 schools.) The Level 3 schools (11 schools in all), by contrast, offered significant structural supports for low achieving students – however, it is significant to note that the CAHSEE alone did not, according to principals, prompt these deeper structural changes. Below the categories are discussed in more detail.

### *Level 1 Support: Before and After School Remediation*

The 17 high schools in this category changed virtually nothing of their school schedule or organizational structure in response to the CAHSEE requirement. The main response of these schools was the addition of CAHSEE review classes before and after school for the 11th and 12th graders who had failed one or both sections of the exam.

According to principals, these ad-hoc before and after school CAHSEE review courses were made mandatory for all students who failed the exam. Yet, the most common observation among principals who instituted this strategy was that student turnout for these CAHSEE preparation courses was disappointingly low. Several principals, however, did acknowledge that students who were already doing poorly in school and had already become academically disengaged were not likely to show up for after school classes. As principal of McKinley High School, a MPR school in this category (with a pass rate of 77% on ELA and 80% on math for 10th graders in spring of 2004), noted of these classes:

the turn out was better for juniors. The juniors are attending; the seniors, it was a ... it was difficult to get the seniors in and – Well, again in looking at the data, many of the students had attendance-some of the students had attendance problems so they're not attending school. So for whatever reason, they've already checked themselves out of the process.

The principal of Coolidge High School, a low performing school with a 10th grade pass rate of just 54% on the ELA and 60% for math in spring of 2004, said that he had considered adding remediation courses during the regular school day for students who had yet to pass the CAHSEE, he decided against it because of his concern that seniors who took these courses would fall too far behind on credits needed for graduation. Yet like the principal of



McKinley, this principal lamented that turnout was quite low for the before and after school prep classes, despite efforts to reach out to the more than 120 seniors who had yet to pass the test.

And now we offer Monday through Thursday CAHSEE preparation for those 120+ seniors first who haven't passed. We've sent the letters home, made the personal contact, the counselors have spoken to them. On our first day, we had two kids show up for English Language Arts, 10 kids show up for Math.

Taken together, the 17 high schools in this category offered few supports for their lowest achieving students, even for their most at-risk students who had already failed the CAHSEE at least once.

While it would be hoped that the very lowest performing high schools in the state would be expected to offer the most intense supports, one-third of the lowest performing schools in the sample (with fewer than 70% passing one or both sections of the test) fell into this category (8 of the 25 total LPR schools in the sample, whose pass rates ranged from 34% to 64% in ELA, and between 40% and 66% in math). These schools' large numbers of failing students were receiving little extra help beyond the after school or before school support classes, which were only offered to them after they had failed the test, at a time when the students were least likely to come.

### *Level 2 Support: Daytime Remediation*

A total of 19 schools decided to offer CAHSEE remediation courses during the regular school day for 11th and 12th graders who had failed the exam in the 10th grade administration. Of the 19 principals in this category, many had said that they had first tried offering the before and after school classes in Level 1, but once they saw that turnout was low, they instead instituted mandatory CAHSEE review courses in the regular schedule. The schools in this category were diverse in terms of initial achievement levels on the 2004 exam, including 12 schools with very LPRs on the CAHSEE (52% of the LPR schools), 4 schools with middle pass rates (29% of MPR schools), and 3 schools with high initial pass rates on the test.

One principal of a school in this category, Lincoln High School (with a 2004 pass rate of 77% for sophomores in ELA and 80% in math), said that this strategy of daytime remediation made it more likely that the school would capture more students in their remediation efforts: "It's not like they have to come to an after-school tutoring program or a Saturday event or-or

some kind of supplementary add-on thing. This is the class they're in! And they have highly-qualified, experienced teachers teaching these classes."

Juniors and seniors enrolled in these classes must give up elective courses in order to have room in their schedule to take the CAHSEE remediation classes. The principal of Cleveland High School – with a pass rate of 93% in ELA and 91% in mathematics for sophomores in 2004 – notes:

And you know, so we changed our curriculum, we had kids give up an elective 2nd semester, and instead, take a ... CAHSEE review English and Math class instead of their elective. "Okay, dude, you're not in PE anymore! You're not taking Art anymore. Dude, you've got to take this class! And you don't have a choice."

Other educators noted that the long turnaround time for results on these exams (10 weeks usually to get results from the state) made it somewhat hard to keep students in these classes, given that many do not want to enroll if they think they have passed the test. As the principal of Ford High School also remarked:

It's always a problem because the – it takes 2 to 3 months to get results back from any state test. And so the kids don't want to take the intervention programs because they feel like, "Well, I think-I think I passed it" or "I think I did well on it this time, so I'm not going to go to the intervention program." Then they find out that they didn't pass it, so ... and then they have a very limited amount of time for the intervention piece.

While double periods of remediation during the day allowed schools to assist more students than the after-school classes, some principals were concerned that these classes may inadvertently push out the least engaged students by taking away their elective courses. As the Fillmore High school principal noted:

You know, when-when you have a kid that hasn't passed the exit exam and, say, he's a senior, for example, and you have-he's got to take, you know, an English class, a Math class, he's got to take Government and Economics, and now all of a sudden, the kid has to take an exit exam Math class or an exit exam English class. Well, you just basically knocked that kid out of any vocational program they could go into. I mean, you know, the-the elective choices for kids are-are severely limited by ... because you know, we want them to pass, we want to try and give them the instruction that they need to get it done, and so we really have limited choices for students and I-I think that that's-I think sometimes people look at and say, "Well ... you know, well, it's okay if they don't get an elective or whatever," but I think when you-you look at what is it that keeps the kid around, what is it that keeps him coming back, but more than that, you know, that elective may not be ... that elective may be ... the-the opportunity for that kid to learn a skill that they can go out into the world and contribute with.

Other principals, however, did not think the removal of an elective was necessarily a bad thing, because from their point of view, the priority was

that the students first learn basic academic skills. Pierce High School principal (MPR) observed:

When it came to the master schedule, you know, it-it-it affects – a senior can't take Art, for example, because they have to still take the mandatory requirements but they have to take this CAHSEE review, so it begins to change the dynamic of what you offered on your campus. I don't necessarily think that's bad. I could have had a senior that would take English, Government, and maybe a Science, and then an elective – a P.E. and a T.A. Well, that's a child that's not really being very well-prepared and maybe really shouldn't be getting a diploma and have these skills that need to be shored up. But they're now taking a CAHSEE review and that review class is going back to those, you know, eighth grade essential standards that they should have mastered then, and now we're helping them to be successful with that.

While the 19 schools in the Level 2 category of support offer somewhat more substantive support than the before and after school tutorial classes, because these classes are only added into the schedule after students fail, these organizational changes are still fairly marginal. These changes constitute essentially a post-hoc remedial effort, an intervention only after students have failed.

### *Level 3 Support: Focus on Prevention and Support*

The greatest level of organizational support offered by schools in this study consisted of both Level 2 remediation (i.e., the addition of CAHSEE remediation courses during the day for 11th and 12th graders who had not yet passed one or both test sections), as well as *preventative* support in the 9th and 10th grades for students who were academically at-risk. This preventative support courses consisted of double periods of content support in either English Language Arts or mathematics for both 9th and 10th graders who were identified as academically low-scoring (based on incoming 8th grade CST scores, teacher recommendation, or 9th grade CST scores). This category consisted of 11 high schools, 5 of which were middle pass rate schools, 1 of which was an HPR, and 5 of which were LPRs.

While the 11 schools in this category offered the most built-in support time for students at all grade levels, not all of these high schools built in the 9th/10th grade support classes in response to the CAHSEE requirement alone. A number of these support classes had been instituted long before CAHSEE was in place – some were instituted in response to California's existing accountability Academic Performance Index (API) system, others instituted these classes before the API system was in place. Thus, the

greatest level of supports offered to students were not, in most cases, prompted by the CAHSEE alone, but rather had already been instituted in response to earlier accountability measures or concerns about student achievement.

What distinguished schools in this category from one another was the range of underlying goals and philosophies undergirding their 9th and 10th grade support courses: some of the principals described these double periods of support as more remedial in nature – that is, intended on giving students basic skills that they were lacking in order to help students reach basic levels of proficiency in the subject area. Other principals described these classes as a means to provide academic support for lower achieving students so they could master more rigorous, college-preparatory content. Most of the schools in the latter category were guided by the Advancement Via Individual Determination (AVID) program philosophy, which provides supports for lower achieving youth to master challenging college-preparatory curriculum. Orange County’s Pierce High School, for example, modeled their double periods of support for low achieving 9th graders after AVID programs:

Most of our students come to us performing well below grade level in reading and on the CSTs, they’re 1s and 2s, far below proficient which the state has mandated. So we have offered a support class – it was really modeled after AVID. They take it as a support class, a second class. It’s backed up with their English class, so they take an English teacher and their support class back-to-back. And we put a focus on reading and writing. And we do enrichment in that, too.

Eisenhower High School, in San Diego County, took the dramatic step, years before CAHSEE was in place, of de-tracking all 9th and 10th grade classes by eliminating honors courses and remedial courses in all subject areas. This racially diverse school is located in a district with a strong commitment to the AVID philosophy, with programs in middle schools as well. The principal observed:

We have what we call an algebra and geometry, block algebra and geometry. So essentially our freshmen and sophomore will take 2 hours. And one hour is the standards, the actual geometry or algebra class. The other hour is a ... is where the teacher assess the students, and the students in there are struggling math students, assess the students, where their holes are in the lower level math, and it gives the teacher the time to teach those, so then they can teach the standards of geometry or algebra. You know, so many times, over the years, math teachers have said, you know, “I can’t teach algebra,” or “I can’t teach geometry” because kids don’t know this.” Well by putting them in the block, and it’s not a homework time or anything else, I make my teachers who teach the block geometry and block algebra, they have to assess the kids right at the beginning. And they have to identify, these are their areas of weakness that are going to

keep them from being successful in algebra or geometry. And then they have to come up with lesson plans to teach those, to fill in those holes.

For the majority of principals interviewed, then, the CAHSEE did not prompt any significant change in school organization, even for those schools with the lowest performance levels in the state. Most schools (36 of the 47 in the study) offered simply ad-hoc support courses either after school or during the school day, but did not engage in any significant re-tooling of their overall school structure. This includes 20 of the 25 (or 80%) of the lowest performing schools in the sample. Of the 11 schools that did offer built in support for students in the 9th and 10th grade, many did not institute this support in response to the CAHSEE alone.

It is important to note that school principals seemed to have a great deal of autonomy (as self-reported) in deciding the changes to institute in their schools. While principals said that districts certainly did mandate particular programs and curricula with regard to the CAHSEE, these mandates almost universally related to the before and after school programs, Saturday courses, and the “crash course” programs. While principals may have overstated their autonomy with respect to school-level decisions in the interviews, it should be noted that, with respect to this issue, there was corroboration across multiple high schools sites within similar districts (indeed, 26 high schools within this study were sampled from within six districts). Thus, from this data it appears that district mandates related largely to the implementation of ad-hoc support courses and curricular packages (and the resources for these packages), and that it was largely up to the principals to decide “what to do” in the context of the regular school day.

## **INSTRUCTIONAL CHANGE: NARROWED CURRICULUM FOR THE LOWEST ACHIEVING STUDENTS**

One of the goals of the CAHSEE requirement was to push schools to align their curriculum to the California content standards, and to prompt schools to address basic skills in areas that students were lacking. The vast majority of principals interviewed for this study said that they were well into aligning their curriculum to the state content standards by the 2004 Spring administration for sophomores in the Class of 2006 (the first administration for the first class to be held to the requirement). Principals did say that this

alignment was further reinforced by the CAHSEE, with its high stakes for students and its major role in NCLB accountability. As the principal of Truman High School in San Diego County (a high performing school with greater than 90% pass rate on both tests) notes, the CAHSEE has “required us to be a more deliberate in our decisions about what we teach and how we deliver instruction,” and has required them to “work harder to make sure that kids are getting those essential standards.”

While standards alignment was the goal of the CAHSEE, a majority of principals (24 of the 47 interviewed) say that the CAHSEE’s influence has extended beyond mere standards alignment, and that the test itself has demanded a share of classroom time. Some of these time demands are small: some schools (12 of the 47) instituted warm-up reviews and brief review activities for the CAHSEE as part of the weekly routine. The principals of another 16 schools said that their teachers were changing the curriculum to directly address the tested material, and of these 16, a total of 5 principals said their teachers were reducing their focus on literature and novels to adjust to their reading and writing assignments to align with the CAHSEE exam questions. Test-prep strategies were also instituted as part of the regular curriculum in 6 of the 47 schools in the study.

### *Reviews and Warm Ups for CAHSEE*

The most minor adjustment made in classroom instruction in response to the CAHSEE was the institution of warm-up reviews in the regular 9th and 10th grade mathematics and English Language Arts courses, which happened in 12 of the 47 schools, according to principal interviews. These reviews were intended to cover 6th and 7th grade standards that many of the students had not seen since middle school. While many high schools used the CAHSEE books provided by the state for their review questions, some high school teachers have called upon their own faculty to develop test review activities, as in Tyler High School, an MPR school in Los Angeles County:

What-what our English Department Chairperson has done is, he’s prepared lessons that we duplicate and give to teachers of tenth grade English right from the beginning of the year-of the school year. And it’s not optional for the tenth grade teachers; they have to go over those lessons every Monday and Wednesday with every one of their tenth grade classes. And now we’ve got all of the English classes reviewing lessons in preparation for the ... the-the state testing in May, the CAT6.

While the CAHSEE exam itself took up some amount of class time in these 12 schools due to the warm up and review activities, it is difficult to know how much class time is actually involved with these activities, and whether these activities lead to a decreased coverage of other grade-level material.

### *Tailored Curriculum to Tested Skills*

Another 12 high school principals said that their teachers had responded to the CAHSEE requirement by tailoring their general ELA and mathematics curriculum to CAHSEE tested materials. This occurred in 9 (36% of the 25) LPR schools and in the low track classes of 1 MPR school (of 14 total) and 2 (of the 8 total) HPR schools.

The principal of Jackson High School, a LPR school in Los Angeles County, talked about how pleased he was that his English Department was tailoring its curriculum to the CAHSEE material:

I gave a letter to my English department thanking them just the day before yesterday 'cause I went in and looked at a lot of our English classes and I did see CAHSEE was on their agenda and some of them had incorporated it into their curriculum ... [Both the test-taking] strategies and also some of the areas that CAHSEE's hitting in particular, they had already interwoven it within their curriculum. And it was in every single English classroom that we visited and I was here with a district team and that-that was great because the district saw [that] we had taken the test and incorporated it into what we were doing in our daily lessons. The Math department has done that even more exclusively. They have designed lessons around their curriculum without extra pay, specific teachers have taken it upon themselves to redesign lessons to incorporate these automatically as the teachers are using them in their classes for-in the Algebra classes and the Geometry classes, any classes that any student is going to be taking. And that includes our 9th graders, by the way, even though the 9th graders are not having to take it. Because they saw the need, they saw the importance, they saw the high stakes of it and so they've taken it upon themselves to do that.

This tailoring of the curriculum to CAHSEE tested material does raise some serious questions about access to curriculum for the lowest-achieving students. For example, the principal of Fillmore High School, a LPR school in Los Angeles County, notes that the embedding of CAHSEE strategies in the low-track courses would never happen in the higher-level classes:

What we do with some of our classes where we may have students that we think-it may be a bit problematic about passing it is we will stress prior to the test some of those skill levels within English Language Arts and Math. Would we do that in an SAS class, School for Advanced [Studies] and Honors class[es]? No. But for some of the other classes, we provide materials for the teachers. Now, what we ask them to do is they have

to embed this – not that they have a lot of time, but they have to embed this in their curriculum delivery as they work with the students prior to the testing, so it's truly like a test prep. We're not teaching to the test, we're not giving questions, certainly, but we're working with those skill areas that we feel are important for our students.

Some would argue that the lower achieving students are benefiting from this increased focus on CAHSEE basics, but this does raise questions about whether CAHSEE further exacerbates what might be called “curriculum gaps,” or inequities in access to rigorous curriculum, between students in the lower- and higher-level courses in schools with tracking structures – or between schools with all low achieving students and schools with higher achieving students (Oakes, Gamoran, & Page, 1992).

### *Reduced Focus on Literature*

The most significant changes made to the regular curriculum as a result of the CAHSEE was in English Language Arts, as many principals said the CAHSEE has led their teachers to reduce their focus on literature. This change was noted by five of the principals interviewed (one of the MPR schools – not mentioned above, and four of the LPR schools mentioned above).

The principals in this study however were not critical of this change (which is not surprising, since they raised the change in the interview). Rather, many saw this change as a necessary trade off to ensure students were graduating with basic skills. As the Fillmore High School Principal (a middle pass rate school in Kern County) notes, his teachers have reluctantly made a necessary shift away from literature to focusing more on basic skills and test-taking strategies:

I mean, you know, there's much more emphasis on grammar and writing and ... you know, sentence structure and all of those, you know, all of those things that ... that really we didn't really work a lot on before. And-and that's probably been, I think, the biggest adjustment, is for English teachers to come to terms with ... with that. [laughs] You know, that's been a huge adjustment for them to make because it's not all about ... Catcher in the Rye or what, you know, what-whatever it may be, Grapes of Wrath. It's about, hey, you know, the kids, you know, they have to learn how to read, reading comprehension, you know, how do you assess that? Doing-working on analogies, you know, teaching kids how to take tests that utilize analogies; I mean, those are all just, like, huge undertakings and that ... that they have to-that they have to work at.

Some principals note that this reduced focus on literature is a necessary change for schools that are very low performing. For example, the principal



of Arthur High School, a low achieving school in Los Angeles County, believes that the reduced focus on literature is good for schools like his:

In English, I think the-the-the change from the text-based thing, like “We’re going to read—we’re going to read *The Scarlet Letter* and wallow around in it however we want.” [Now] “We’re going to do these standards and we may only read part of the book.” It’s probably a good thing for an under-performing school. I’m not sure that’s a good decision for a higher-performing school... But I think that’s good. The teachers are becoming more aware of it. I mean, I think it helps them understand how they have to deconstruct their lessons. So I think a lot of good things have happened.

This reduced focus on literature and increased focus on basic skills in the lower performing schools (and in classes serving the lowest achieving students) again raises serious questions about a curriculum gap between schools serving higher achieving and lower achieving students (and within schools with both higher and lower achieving students as well). Some would say that this increased focus on CAHSEE content in these classes is beneficial to students who had yet to master these basic skills, but what remains unclear is the degree to which students are being taught a more remedial curriculum vis-à-vis their higher achieving peers – as opposed to receiving supports to master the rigorous and engaging grade-level curriculum (Burris & Welner, 2005).

### *Increased “Test Prep” Activities*

A handful of principals in this study (six total) said that their classroom teachers taught test-taking strategies to their students as a result of the CAHSEE in the general curriculum. Of the six principals (which comprised 13% of the sample), five were LPR schools (of the 25 total) and one was an MPR school. According to principals, these test-taking strategies were being taught in classes in response to *both* the CAHSEE and the CSTs.

One school teaching such strategies, Lincoln high school (an MPR school in Orange County) says that they adopted these strategies to give students a better understanding about how to take tests, to eliminate any errors caused by a misunderstanding of test format:

We have—we have a school-wide emphasis and understanding that state wide high-stakes testing is important. So we – two years ago, we in-serviced all of our core subject areas: Math, English, Science, and Social Studies teachers so that every single one of the teachers were trained in Princeton Review testing strategies and we purchased all of the materials for test prep weekly activities in the classroom. So for the past two years and

continuing into this year, we've been doing that. It's integrated into the curriculum and instruction.

Thus, only a handful of principals responded to the accountability requirements imposed by the CAHSEE and the CSTs by instituting test-taking strategies as part of the regular curriculum. Again, from principal interviews it is difficult to discern how much these test-prep activities figure into the regular curriculum – and whether these activities mean that less time is available to cover other content material. More in-depth investigation is needed to better understand this phenomenon.

### *CAHSEE Remediation Classes*

While few principals instituted test-prep strategies in their regular ELA and mathematics curriculum, test-prep activities figured a large role in virtually all of the CAHSEE remediation classes (either during the regular school day or before and after school) that were offered for 11th and 12th graders who had yet to pass the exam. These CAHSEE remediation classes, typically named “CAHSEE Preparation” or “CAHSEE Essentials,” typically included a mixture of both test-prep strategies and CAHSEE content focus in the curriculum.

A large majority of high schools turned to outside test-preparation vendors like Kaplan to provide the curriculum to these classes. One of these schools was Washington High school, a high performing school in Orange County, which noted that Kaplan provided textbooks and workbooks:

We also have CAHSEE remediation classes for students – now, these are for kids after their sophomore years who haven't been successful. And so we have CAHSEE remediation classes that focus on looking at specific questions, strategies for answering questions, test anxiety, all of those factors that play into taking the standardized test. And we use a series put out by Kaplan which is an actual textbook-type program and a workbook.

Some schools instituted academic “boot camps” for seniors who had failed the test. These boot camps, made mandatory usually by the schools or districts, required students to be pulled out of all of their regular courses for 1 or 2 weeks and given intensive CAHSEE – geared instruction. Other schools became more strategic with their remediation efforts, like Fillmore High School, a middle performing school in Kern County. The principal observed that his teachers recognized the difficulty of students mastering algebra in time to pass the CAHSEE, so that they decided in some cases to

focus on the pre-algebra skills so that students could get enough of those questions correct to get a passing score on the CAHSEE. He notes:

I mean, especially with our kind of down to the wire pull-out preparation, we've tried not to focus a lot on the Algebra component as much as we have the other areas just because they can pass without really having all of the Algebra and to try and get over that hurdle in a couple of weeks is too big of a hurdle for the kids to get over.

According to principals, then, the CAHSEE remediation courses offered to 11th and 12th graders in most schools are not intended to ensure students really understand the content, but more of an effort to get the students a passing score on the test. This raises questions about whether the CAHSEE is really prompting the kinds of content mastery that was desired by policymakers when they instituted the test.

### *The Relationship between Structural Support and Curricular Change*

To fully understand the impact of the CAHSEE on high schools, it is important to examine the types of curricular and instructional changes that are made by schools in the presence or absence of other organizational support structures. In other words, it is not possible to fully understand the impact of organizational changes (or lack thereof) without understanding how a school's curriculum and instruction changed along with it.

Examining both dimensions of change together, it becomes apparent that where low-achieving students were getting the least structural supports (in terms of 9th and 10th grade academic enrichment classes that are offered in addition to the core curriculum), the curriculum was also, according to principals, the most CAHSEE-driven. These schools, which only offered before and after school CAHSEE prep for 11th and 12th graders, were also more likely, according to principals, to tailor their regular curriculum in the 9th and 10th grade in particular to CAHSEE content. By contrast, principals of schools that offered relatively more academic supports for low achieving students (by offering academic support courses in 9th and 10th grade in addition to daytime CAHSEE preparation courses in 11th and 12th grades) tended to describe little change to their regular 9th and 10th grade curriculum; rather these principals sought to keep their curriculum focused on grade-level standards while at the same time providing supports for low achieving students to catch up at other times during the school day.

Of the 17 schools in the “Level 1” category of organizational change (which offered only before and after school CAHSEE preparation for students who have failed the CAHSEE), a total of nine principals said they changed their regular 9th and 10th grade curriculum to focus on specifically on tested content. Of these nine schools, six were low performing (LPR) schools, and three were higher achieving schools (one was an MPR school and two were HPR schools) where principals noted that the low track courses had changed to focus on CAHSEE preparation. Of these low-support “Level 1” schools, furthermore, three principals said that their ELA teachers had also reduced the focus on literature (two of which were LPR schools and one was an MPR school).

Curricular changes were apparent but less significant in the schools in the Level 2 category of organizational change – those schools that offered CAHSEE remediation courses during the school day for 11th and 12th graders who had yet to pass the test. Half of principals in this category (nine total) said that their teachers did review CAHSEE material prior to the test through warm ups or review activities. Two of the principals of schools in this category said their teachers directly tailored their curriculum to CAHSEE-tested content by reducing their focus on in-depth literature in favor of CAHSEE-tested short passages, and an additional four principals said that their school had instituted test-prep strategies as a part of their regular curriculum.

By contrast, the principals of the 11 “Level 3” schools that offered the greatest structural supports for their low achieving students (in terms of time outside of the regular curriculum to catch up on skills) described less significant changes in their regular curriculum in response to the CAHSEE test itself. Just one principal (of an LPR school) said that the school’s teachers reduced their focus on literature in the context of the regular ELA curriculum, and no other principals described any other changes to their regular curriculum related to the CAHSEE exam itself. Just over half of these principals (6 of the 11) described offering warm ups in the content of their regular 9th and 10th grade curriculum.

In sum, students who were getting the least structural supports were also, according to principals, getting the most CAHSEE-driven curriculum. Instead of adopting longer-term preventative strategies to address underachievement for their most academically at-risk students, the lowest performing schools reacted with a “too little, too late” strategy that consisted often of test-prep and remediation in the 11th and 12th grades. An obvious danger of waiting until 11th and 12th grade to intervene is that students who are academically at-risk of failing the CAHSEE are not receiving support until they have fallen far behind.

*Implications for Policy and Practice*

The California High School Exit Exam was created to pursue important policy objectives: the CAHSEE was intended to prompt educators to work harder to ensure that students acquired basic skills; students were hoped to take greater responsibility for their learning.

We have seen how the CAHSEE did accomplish one of its aims: schools did become more focused on their lower achieving students as a result of this requirement. However, the ways in which they reorganized curriculum and instruction *for* those students raises serious questions about the ultimate benefit of this policy.

Few schools offered serious preventative support for students who were academically low achieving early on in their educational careers. Rather, most of the supports that high schools offered were not available to students until after they had failed the test. Many of these support classes were only made available to students after school, at the time when the least engaged students were least likely to take them. Furthermore, many of these support classes were not necessarily designed to ensure students master the content – many were as focused on test-prep strategies as they were on the content area, according to principals.

Many schools changed their core curriculum in response to the CAHSEE, but these changes described by principals typically involved tailoring the curriculum to focus more on tested material. These changes were particularly notable in English Language Arts, where many principals reported a reduced focus on literature, particularly for the lowest achieving students. This raised the serious question of “curriculum gaps” between higher and lower achieving students within schools, and *between* schools serving higher and lower achieving students.

When looking at organizational and instructional change together – how curriculum changed in schools that provided, or did not provide, built in time for students to receive academic support – interviews revealed that the schools that proved best able to support students and design programs that were truly supportive (rather than remedial) in nature, were those schools that seemed to be high-functioning already, and which had a coherent strategy for increasing the achievement of their lowest performing students (Level 3 schools). By contrast, schools that were not doing well did not fundamentally change their school organization or their curriculum as a result of the CAHSEE requirement; these schools instead reacted with a “too little, too late” strategy that consisted often of test-prep and remediation in the 11th and 12th grades.

Ultimately policymakers may not be creating the kinds of changes they were hoping for when they instituted this policy. This is in part because this policy came without any state resources or investments in capacity in terms of leadership knowledge at the district or school level, instructional knowledge at the school and classroom level, or in terms of fiscal capacity to acquire the necessary resources that are required to institute change (Elmore, 2003; Elmore, Peterson, & McCarthey, 1996; Fullan, 2001; Spillane & Seashore Louis, 2002).

Although the state did grant school districts a one-time allocation of funds for providing support to seniors in the Class of 2006 who had yet to pass one or both portions of the exam in the fall of 2005 (which was beyond the period of data collection for this project),<sup>3</sup> this one-time allocation of remediation funds restricted districts' use of the funds to programs and supports for seniors who had yet to pass the CAHSEE. As such, these supports only aided districts' efforts to institute post-hoc support rather than adopt preventative strategies. It is unclear, furthermore, whether these remedial supports will remain in place over time: the Center on Education Policy has found that other states with exit exams have cut similar "emergency" remedial funding after their exams had been in place a number of years (CEP, 2006).

California's exit examination policy, therefore, was instituted without any serious investments at state level in cultivating the ability of schools, as institutions, to better serve their most at-risk student populations. As Richard Elmore (2002) notes, school change is entirely dependent upon the capacity of those charged with implementing that change:

Low-performing schools, and the people who work in them, don't know what to do. If they did, they would be doing it already. You can't improve a school's performance, or the performance of any teacher or student in it, without increasing the investment in teachers' knowledge, pedagogical skills, and understanding of students. This work can be influenced by an external accountability system, but it cannot be done by that system (pp. 33-34).

The state can foster its goal of increasing student achievement by providing schools both with the resources required by the test (in terms of off-setting the monetary and personnel costs born by schools) as well as the resources required for basic school operations (qualified teachers, adequate materials); as well as technical assistance and guidance to help schools identify longer-term reform strategies and instructional strategies for students – rather than interventions that amount to “too little, too late.”

These interventions, as the principals themselves noted, were needed far earlier in the K-12 pipeline. By the time students at risk of not passing the

CAHSEE reached high school, they were often several grade levels behind. As research has consistently shown, these gaps begin even before students enter kindergarten (Lee & Burkham, 2002). Those schools that wait until 11th or 12th grades to do any serious intervention with students end up catching students when they have fallen far behind and are at the greatest risk of becoming discouraged by the CAHSEE, at an extremely risky juncture in their educational career.

Another unintended consequence of this reform is the growing “curricular gaps” that principals had described between the highest and lowest performing students – both within and between schools. If the goal of this policy is to prepare students for the 21st century economy, policymakers need to be aware that some of the responses may be counter to the goals of the policy, as many schools seem to be gearing their curriculum for low achieving students towards basic skills.

Overall, these findings suggest that the exit exam is not spurring the kinds of systemic improvements California’s policymakers were hoping for when they instituted this policy. These disappointing results stem in part from the lack of state investment in capacity building at the district and school level (Elmore, 2004). This suggests that states cannot expect improvements to flow directly from a test mandate alone; without investments in resources and knowledge, schools will respond with incremental, ad-hoc changes that fail to live up to the underlying goals of the policy itself.

## NOTES

1. It should be noted that CAHSEE scores from the Spring of 2002 were used for NCLB purposes, as 2002 was the baseline year for NCLB. This 2002 test, however, was an older version of the exam, significantly different from the exam used from Spring 2004 onward.

2. The names of the high schools in this chapter have been changed to protect the confidentiality of the principals.

3. This remedial funding, made possible by AB 128 (Chapter 234, Statutes of 2005), was funded through a one-time apportionment of \$600 per senior who had failed one or both portions of the CAHSEE (<http://www.cde.ca.gov/fg/fo/r19/cahseeii05apptlter1.asp>).

## ACKNOWLEDGMENTS

This research was generously supported by a grant from the Spencer Foundation (Grant #200600108). All conclusions and analyses are the

author's own. A preliminary version of this paper was presented at the 2007 Annual Meeting of the American Educational Research Association in Chicago, IL. This paper benefited from the insightful feedback of Professor Jeannie Oakes of UCLA, and from the helpful suggestions of RSE's anonymous reviewers.

## REFERENCES

- Achieve, Inc. (2005). *Action agenda*. Retrieved on 8/12/05 from [http://www.achieve.org/achieve.nsf/ActionAgenda\\_Expectations?openform](http://www.achieve.org/achieve.nsf/ActionAgenda_Expectations?openform)
- Amerin, A. L., & Berliner, D. C. (2002). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in states with high school graduation exams*. Tempe, AZ: Arizona State University, Educational Policy Studies Research Unit.
- Beyer, L. E. (1985). Educational reform: The political roots of a nation at risk. *Curriculum Inquiry*, 15(1), 37–56.
- Burris, C., & Welner, K. (2005). Closing the achievement gap by detracking. *Phi Delta Kappan*. Retrieved on 4/11/05 from [http://www.pdkintl.org/kappan/k\\_v86/k0504bur.htm](http://www.pdkintl.org/kappan/k_v86/k0504bur.htm)
- California Department of Education (CDE). (2006). News release: Schools Chief Jack O'Connell holds firm on high school exit exam; recommends options for struggling students. Accessed on 3/14/06 from <http://www.cde.ca.gov/nr/ne/yr06/yr06rel4.asp>
- Carnoy, M., Elmore, R. F., & Siskin, L. S. (Eds). (2003). *The new accountability: High schools and high-stakes testing*. New York: Routledge and Falmer.
- Center on Education Policy (CEP). (2006). *State high school exit exams; a challenging year*. Washington, DC: Center on Education Policy.
- Center on Education Policy. (2007). *State high school exit exams: Working to raise test scores*. Washington, DC: Center on Education Policy.
- Darling-Hammond, L., Rustique-Forrester, E., & Pecheone, R. (2005). *Multiple measures approaches to high school graduation*. Palo Alto, CA: Stanford School Redesign Network.
- Dee, T. S., & Jacob, B. A. (2006). *Do high school exit exams influence educational attainment or labor market performance?* Working Paper 12199. Washington, DC: NBER. Accessed on 11/4/06 from <http://www.nber.org/papers/w12199>
- Diamond, J. B., & Spillane, J. P. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record*, 106(6), 1145–1176.
- Elmore, R. (2002). Unwarranted intrusion. *Education Next*, 2(1), 31–35.
- Elmore, R. F. (2003). Accountability and capacity. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *In the new accountability: High schools and high stakes testing* (pp. 195–209). New York: Routledge and Falmer.
- Elmore, R. F., Peterson, P. L., & McCarthy, S. J. (1996). *Restructuring in the classroom: Teaching, learning & school organization*. San Francisco: Josey-Bass.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Gayler, K. (2005). *How have high school exit exams changed our schools? Some perspectives from Virginia and Maryland*. Washington, DC: Center on Education Policy.



- Gosslein, P.G. (2006, March 6). That good education might not be enough; American workers at all levels are vulnerable to outsourcing, experts say, posing a challenge to the assumption that more schooling is the answer. *Los Angeles Times*, p. A1. Accessed [11/15/06] via: <http://proquest.umi.com/pqdweb?did=997708001&sid=1&Fmt=3&clientId=1564&RQT=309&VName=PQD>
- Greene, J. P., & Winters, M. (2004). *Pushed out or pulled up? Exit exams and dropout rates in public high schools*. New York: Center for Civic Innovation at the Manhattan Institute.
- Hamilton, L. S., & Koretz, D. M. (2002). Chapter 2: Tests and their uses in test-based accountability systems. In: L. S. Hamilton, B. M. Stecher & S. P. Klein (Eds), *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.
- Heubert, J. P., & Hauser, R. M. (Eds). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- HumRRO. (2004). *California High School Exit Examination (CAHSEE): Year 4 evaluation report*. Alexandria, VA: HumRRO.
- HumRRO. (2006). *Independent evaluation of the California High School Exit Examination (CAHSEE): 2006 Evaluation report*. Alexandria, VA: HumRRO.
- Lee, V., & Burkham, D. T. (2002). *Inequality at the starting gate*. Washington, DC: Economic Policy Institute.
- McNeil, L., & Valenzuela, A. (2001). The harmful impact of the TAAS system of teaching in Texas: Beneath the accountability rhetoric. In: G. Orfield & M. Kornhaber (Eds), *Raising standards or raising barriers?* (pp. 127–150). New York: Century Foundation.
- Mintrop, H. (2003). The limits of sanctions in low-performing schools: A study of Maryland and Kentucky schools on probation. *Education Policy Analysis Archives*, 11(3). Retrieved on 4/24/06 from <http://epaa.asu.edu/epaa/v11n3.html>
- Mishel, L., Bernstein, J., & Allegretto, S. (2006). *The state of working America: 2006/07*. Washington, DC: Economic Policy Institute.
- Muller, C., & Schiller, K. S. (2000). Leveling the playing field? Students' educational attainment and states' performance testing. *Sociology of Education*, 73(3), 196–218.
- Oakes, J., Gamoran, A., & Page, R. N. (1992). Curriculum differentiation, opportunities, outcomes, and meanings. In: P. W. Jackson (Ed.), *Handbook of research on curriculum* (pp. 570–608). Washington, DC: American Educational Research Association.
- Oakes, J., Wells, A., Jones, M., & Datnow, A. (1997). Detracking: The social construction of ability, cultural politics, and resistance to reform. *Teachers' College Record*, 98(3), 482–510.
- O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72(3), 293–327.
- Olson, L. (2006). *Economic trends fuel push to retool schooling*. Accessed on 5/19/06 from <http://www.edweek.org/ew/articles/2006/03/22/28prepare.h25.html>
- Rhoten, D., Carnoy, M., Chabran, M., & Elmore, R. F. (2003). The conditions and characteristics of assessment and accountability. In: M. Carnoy, R. Elmore & L. S. Siskin (Eds), *In the new accountability: High schools and high stakes testing* (pp. 13–54). New York: RoutledgeFalmer.
- Rogers, J., Holme, J. J., & Silver, D. (2005). *More questions than answers: CAHSEE results, opportunity to learn, and the class of 2006*. UCLA: UCLA's Institute for Democracy, Education and Access. Available at <http://www.idea.gseis.ucla.edu/resources/exitexam/index.html>

- Skra, L., Scherurich, J. J., Johnson, J. F., & Koschorek, J. W. (2004). Complex and contested constructions of accountability and educational equity. In: L. Skrla & J. J. Scheurich (Eds), *Educational equity and accountability: Paradigms, policies and politics*. New York: Routledge Falmer.
- Spillane, J., & Seashore Louis, K. S. (2002). School improvement processes and practices: Professional learning for building instructional capacity. In: J. Murphy (Ed.), *Challenges of leadership. Yearbook of the National Society for the Study of Education*. Chicago, IL: University of Chicago.
- Trumbull, M. (2006). US economy's latest output: Better jobs. *Christian Science Monitor*. Retrieved on 12/20/06 from <http://www.csmonitor.com/2006/0411/p01s02-usec.html>
- Tyack, D. (1993). School governance in the United States: Historical puzzles and anomalies. In: J. Hannaway & M. Carnoy (Eds), *Decentralization and school improvement*. San Francisco: Jossey-Bass.

APPENDIX

CAHSEE Performance Spring 2004 Class of 2006			Response		Location and Type of School			
ELA Spring of 2004 10th (%)	Math Spring of 2004 10th (%)	Category of pass rate status	Overall level of support	“Tailored” curriculum?	Test prep?	County	Location	Type of school
98	97	HPR	Level 1	X		San Diego	Suburban	Comprehensive
90	90	HPR	Level 1			Los Angeles	Rural	Comprehensive
91	91	HPR	Level 1			San Diego	Suburban	Comprehensive
91	93	HPR	Level 1	X		Orange county	Suburban	Comprehensive
59	65	LPR	Level 1	X		Alameda	Urban	Comprehensive
60	57	LPR	Level 1	X	X	Los Angeles	Urban	Comprehensive
54	60	LPR	Level 1	X		Sacramento	Urban	Comprehensive
34	40	LPR	Level 1	X		Sacramento	Urban	Charter, small
64	66	LPR	Level 1	X	X	Los Angeles	Urban	Comprehensive
58	54	LPR	Level 1	X		Kern	Suburban	Comprehensive
61	58	LPR	Level 1			Los Angeles	Urban	Comprehensive
34	40	LPR	Level 1			Sacramento	Urban	Charter, small
74	72	MPR	Level 1	X		Los Angeles	Urban	Comprehensive
80	77	MPR	Level 1			Sacramento	Urban	Comprehensive
77	80	MPR	Level 1			Los Angeles	Urban	Comprehensive
70	82	MPR	Level 1			Los Angeles	Suburban	Comprehensive
86	89	MPR	Level 1			San Diego	Urban	Comprehensive
86	88	MPR	Level 2			Sacramento	Suburban	Comprehensive
92	96	HPR	Level 2			Orange	Suburban	Comprehensive
92	92	HPR	Level 2			Orange	Suburban	Comprehensive
58	58	LPR	Level 2	X		Kern	Urban	Comprehensive
49	32	LPR	Level 2		X	Los Angeles	Urban	Comprehensive
55	55	LPR	Level 2		X	Los Angeles	Urban	Comprehensive

APPENDIX. (Continued)

CAHSEE Performance Spring 2004 Class of 2006			Response		Location and Type of School			
ELA Spring of 2004 10th (%)	Math Spring of 2004 10th (%)	Category of pass rate status	Overall level of support	“Tailored” curriculum?	Test prep?	County	Location	Type of school
50	48	LPR	Level 2		X	Los Angeles	Urban	Comprehensive
68	74	LPR	Level 2			Alameda	Urban	Small school
58	45	LPR	Level 2			Alameda	Urban	Small school
49	53	LPR	Level 2			Alameda	Urban	Small school
65	55	LPR	Level 2			Los Angeles	Urban	Comprehensive
64	68	LPR	Level 2			Los Angeles	Urban	Comprehensive
46	50	LPR	Level 2			Kern	Rural	Comprehensive
75	69	LPR	Level 2			San Diego	Urban	Small school
58	57	LPR	Level 2	X		Kern	Suburban	Comprehensive
77	80	MPR	Level 2		X	Orange	Suburban	Comprehensive
75	73	MPR	Level 2			Orange	Urban	Comprehensive
74	70	MPR	Level 2			Sacramento	Suburban	Comprehensive
93	91	HPR	Level 2			San Diego	Suburban	Comprehensive
95	96	HPR	Level 3			Orange	Suburban	Comprehensive
70	68	LPR	Level 3		X	Los Angeles	Urban	Comprehensive
60	64	LPR	Level 3			Sacramento	Urban	Comprehensive
67	63	LPR	Level 3			Los Angeles	Urban	Comprehensive
76	73	MPR	Level 3			Los Angeles	Urban	Comprehensive
82	81	MPR	Level 3			San Diego	Suburban	Comprehensive
69	72	LPR	Level 3			Orange	Suburban	Comprehensive
71	74	MPR	Level 3			Orange	Urban	Comprehensive
83	89	MPR	Level 3			Los Angeles	Suburban	Comprehensive
64	73	LPR	Level 3			Los Angeles	Suburban	Comprehensive
82	85	MPR	Level 3			Sacramento	Suburban	Comprehensive

# CHAPTER 7

## DISTRICT CAPACITY AND ACCOUNTABILITY: PROFESSIONAL DEVELOPMENT AS REFORM TOOL

Soung Bae

### ABSTRACT

*Scholars and reform activists see district-level leaders as key actors in improving teaching and learning. This study examines the efforts of one district that successfully narrowed achievement gaps by largely focusing on teacher professional development. I employ the concepts of physical capital, human capital, and social capital as key ingredients of the process of instructional reform. I highlight the district's role in creating system-wide changes in instruction through investment in developing teachers' knowledge and pedagogical skills.*

The growing scholarship on the role of school districts in improving teaching and learning demonstrates that school district leaders can play a potent role in the school improvement process (Elmore & Burney, 1997; Fullan, 1985; Marsh et al., 2005; Massell, 2000; Spillane, 1996; Spillane & Thompson, 1997; Togneri & Anderson, 2003). These studies highlight the critical links between the central office and the process of improving instruction and boosting achievement. McLaughlin and Talbert (2003)

---

Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability  
Research in Sociology of Education, Volume 16, 189–207  
Copyright © 2008 by Emerald Group Publishing Limited  
All rights of reproduction in any form reserved  
ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16007-4

contend that the district office is a significant agent in advancing or reducing the improvement of teachers' instructional practices through its policy-making and leadership. In addition, [MacIver and Farley's \(2003\)](#) review of the literature on central office studies suggests that the district's role in improving instruction and achievement can be distilled to three components: decision-making about curriculum and instruction, supporting effective pedagogy through professional development, and evaluating results to feed back into instructional practices.

This chapter focuses on the school district's role in supporting teacher learning and development. It elaborates findings previously outlined in our research team's working paper ([Woody, Bae, Park, & Russell, 2006](#)). This chapter digs deeper by situating the findings in a broader discussion of the district's role in building teacher capacity. I present a case study of one district's reform efforts that focused specifically on teacher professional development. This chapter describes and analyzes the district's strategic use of professional development to build capacity towards the goal of improving teaching and learning. In doing so, I illuminate the district's role in creating system-wide changes in instruction through the investment of developing teachers' knowledge and pedagogical skills.

## THEORETICAL GROUNDING

To understand district leader's influential role in building capacity through teacher professional development, I employ the theoretical framework of [Spillane and Thompson \(1997\)](#) who measure the district's ability to support instructional reform via physical capital, defined as financial and material resources in addition to available time and staffing; human capital, defined as individuals' knowledge and skills as well as their commitment and disposition to engage in the reforms; and social capital, defined as professional networks and trusting collegial relations. [Spillane and Thompson's \(1997\)](#) analysis demonstrates the interwoven nature of these forms of capital. They argue that the development of human capital (teacher learning) is critically dependent on the development of social capital (vital connections to sources of knowledge outside the district and norms of collegiality within). Without norms of trust and collaboration, professional learning and development will most likely be limited. And although physical capital is necessary to support ambitious instructional reform, it is not sufficient without the related components of human and social capital.

The need for developing human capital is evidenced by the current emphasis on standards-based instruction and high-stakes accountability reform. These reforms may require teachers to make significant changes to their existing practices in order to meet higher standards (Ball & Cohen, 1999; Little, 1993). This contemporary emphasis raises the bar for effective professional development since it is considered an essential mechanism for deepening teachers' content knowledge and developing their instructional skills (Desimone, Porter, Garet, Yoon, & Birman, 2002). Research on effective professional development models suggests that there are several factors that contribute to teacher learning. The common elements of effective professional development include increasing teachers' subject matter knowledge (Ball & Cohen, 1999; Borko & Putnam, 1995; Hawley & Valli, 1999; Wilson & Berne, 1999), creating communities of practice (Ball & Cohen, 1999; Borko, 2004; Hawley & Valli, 1999; Wilson & Berne, 1999), examining student work (Ball & Cohen, 1999; Borko, 2004; Sykes, 1999), and acknowledging the role of teachers' beliefs (Borko & Putnam, 1995; Hawley & Valli, 1999). Richardson (2003) asserts that the learning opportunities offered to teachers should include the following characteristics: be schoolwide; be sustained and include steady follow up; have administrative support; be adequately funded; and develop buy-in from participants. Further, studies demonstrate that frequent participation in high-quality professional development are associated with teachers' use of innovative instructional practices that promote deep learning and engagement (Cohen & Hill, 2000; Smylie, Allensworth, Greenberg, Harris, & Luppescu, 2001).

Similar to the development of human capital, increasing teachers' social capital is an important component of ambitious instructional reform. Coleman (1988) posits that social capital operates in the structure of relations between and among actors. Social capital facilitates the actions of actors within the structure, without which certain achievements would not be possible. Social capital in the form of professional networks provides teachers with access to resources and opportunities for collaboration and learning. Opportunities to learn via social interactions can enable improvements in instructional practice. Studies have shown that the strength of social capital among teachers has a positive effect on teachers' commitment to progressive beliefs and practices (Bidwell & Yasumoto, 1999) and on the implementation of instructional innovations (Frank, Zhao, & Borman, 2004). Therefore, attention must be paid towards the ways in which school districts enable the cultivation of social capital.

## METHODS

### *District Context*

Central Unified School District is an urban fringe district located near a mid-sized city in California's central valley.<sup>1</sup> The district enrolls over 10,000 students in their elementary, middle, and secondary schools.<sup>2</sup> The district registers an ethnically diverse group of students with 53% identified as Hispanic, 33.8% as White, 5% as Asian, and 3.1% as African-American. Nearly 45% of the student population qualifies for free or reduced-price lunch.

Student achievement in the district has climbed steadily since 2002. Each year, student performance, as reflected in the Academic Performance Index (API) has demonstrated growth in each of the district's significant subgroups.<sup>3</sup> During the 2004–2005 school year, 97% of the district's teaching staff was fully credentialed, compared to the statewide average of 93%. The district employs only 3% of its teachers on emergency credentials or waivers, while 4% of the state's teachers fall into that category. The average years of teaching experience for the district's teacher workforce is 14 years, with an average of 11 years teaching within the district. This is over the state average of 13 years of teaching experience and an average of 10 years teaching in a single district. These statistics suggest that the teacher workforce in Central is a relatively veteran faculty with high retention rates.

### *Data Collection*

A team of researchers conducted site visits and data collection for the study during the 2005–2006 school year. The data were gathered through semi-structured interviews with key district and school personnel and observational fieldwork. The purpose of the interviews was focused on gaining an understanding of the district's instructional reform strategy. As such, the interview questions centered on how the district develops and implements opportunities for teacher learning; provides access to material and financial resources; and structures opportunities for teacher collaboration and interaction. The data included interviews with the superintendent, the assistant superintendent, the director of curriculum and instruction, and a program specialist. Focus group interviews were conducted with the coaching staff, the Beginning Teacher Support and Assessment staff, and teachers and the principal from one elementary school. The elementary



school, which the research team visited and conducted the focus group teacher interview, was chosen by the director of curriculum and instruction. The team viewed the focus group as typical of the district's teachers since their years of teaching experience in the district ranged from 8 to 25 years and the average years of teaching experience of the focus group teachers was 18 years.

In total, the research team formally interviewed 13 district office personnel and 7 school personnel. All of the interviews lasted from one to two hours and were tape-recorded and transcribed. A district-wide professional development session for first-grade teachers was observed and extensive field notes were prepared. During this time, the research team spoke informally with the 18 teachers about the professional development session and their experiences teaching in the district. Finally, central office documents such as district improvement plans and mission statements were collected.

### *Data Analysis*

Our review of observational, interview, and archival data sought to understand how the district had built capacity through its strategic use of professional development. We know that the district had successfully narrowed achievement gaps. Did its concentrated focus on professional development contribute to their progress? The interview and observational data were analyzed via multiple passes. The first pass involved reading the interviews and field notes to get a sense of the data and emergent themes. Next, each interview and observation field note was coded using NVivo, a program designed for qualitative data analysis. The coding categories were created a priori and included physical, social, and human capital. The physical capital code captured mentions of material and financial resources such as funding, teaching materials, meeting time, and new leadership. The social capital code encompassed comments about administrative support, teacher learning communities, feeling heard by others, and being invited into classrooms. The human capital code included mentions of district-wide professional development activities, teacher site visits, and the coaching model. Other categories such as history of staff development and student needs emerged directly from the data. Afterward, I moved to a more explanatory level that involved creating thematic matrices from patterns emerging from the data (Miles & Huberman, 1994).

## FINDINGS

Central's reform strategy attended to the development of human and social capital through their strategic use of physical capital. The development of human capital began with the establishment of a clear vision for the improvement of teaching and learning in the district. This vision led the district to create a new professional development model that was coordinated and district-wide. The development of social capital focused on the new coaching model that was instituted by the central office. The coaching model provided teachers with the technical and social support needed to make improvements in their learning and teaching of new instructional approaches.

### *Human Capital: Establishing a Coherent Focus on Teaching and Learning*

The current superintendent assumed leadership of the district in 2001. One of his top priorities became the development of a strategic plan to raise student achievement and meet the needs of all the students in the district. District leaders attribute rising student achievement to major shifts in the priorities of the district. The director of curriculum and instruction recalled, "One of the first things the superintendent did was he convened a task force that first year and they wrote a strategic plan ... it prioritized the goals of the district. And I think there are 22 goals now, but the main thing is they have faithfully from the board down, followed that priority list." Multiple references to the strategic plan came through many of the interviews with district administrators and teachers. One teacher described how the strategic plan "was like a blaring trumpet" extolling the district's instructional focus.

The development of the district's strategic plan led to prioritizing the district's goals. As a result, the number one priority of the district became increasing student achievement in reading, language arts, and writing. The choice to prioritize reading achievement was based on the perspective that reading is the building block on which all other subjects rest. An assistant superintendent explained, "Reading is definitely the basis for student success. If they can't read, they can't do math, they can't do science, they can't do social studies." To work in this district means one is well aware of the literacy priority. When asked what was going on in the district that was making the students' achievement scores go up, one teacher replied, "Good teaching practices and a focus on literacy." Another teacher stated simply, "Literacy, literacy, literacy." Another teacher summed it up by

saying, “The goal of literacy is overall. Everyone knows. Everyone knows what that goal is. We know that that is the goal the district has been pushing.”

### *District-Wide Professional Development*

The district’s conception of teacher professional development has changed since the development of the district’s strategic plan. Prior to planning effort, Central offered a menu of learning options or what the superintendent described as “smorgasbord” training. The professional development opportunities were provided by the district or by outside educational organizations and teachers could decide whether or not they attended. This created a “have or have not” scenario stated the director of curriculum and instruction. “Before, it was some people went to a lot and some people went to just a few things.” As a result, the Central teachers reported that the professional development seemed to lack coherence or purpose. One teacher explained saying, “Before, we had lots of professional development but there was somebody from here that did their thing and somebody from there that did their thing. There were a lot of outside people that go away and you never see again. Nobody really knows what they said.” Another teacher agreed, “There was a little bit here and a little bit there and nothing was mandated one way or the other that you had to do it that way.” Consequently, learning opportunities varied considerably throughout the district.

In an effort to provide coherence to the district’s professional development and get all the teachers on the same page, the superintendent reported that a lot of hard work went into bringing the teachers together, literally. He felt that it was vital to have every teacher present for staff development in order to increase “the power of the staff development getting into each classroom.” Thus, he embedded staff development days into the teacher contracts and developed a coordinated approach to teacher learning. An assistant superintendent echoed the sentiment by stating, “I think the current philosophy within the district, that is embraced highly, is that we all need to grow and we all need to learn. We all have areas in which we need to improve and that we can’t expect kids to make progress if we don’t ourselves engage in ongoing learning. So staff development of some sort is a key focus in the district at all times.”

Currently, the professional development opportunities for Central’s teachers are developed “in house.” The director of curriculum and

instruction and the coaching staff, who are teachers on special assignment, create the learning agendas based on analyzing students' progress on state and district tests, familiarizing teachers with newly adopted curricula, as well as surveying the teachers' needs. In addition, the professional development is provided district-wide. The research team attended a two-day professional development on differentiated instruction for first-grade teachers. For this, the district hired substitutes for half of the first-grade teachers to attend the two-day learning workshop one week, and then provided the same learning opportunities for the other half of the first-grade teachers the following week. The professional development was developed and led by two teachers on special assignment: one had an English Language Development background and the other had a reading specialist background. The professional development leaders were observed providing the teachers with current research on differentiation, allowing teachers time to talk among their colleagues and reflect individually about their understanding and use of differentiation strategies, and providing time to plan collaboratively among school teams various ways to differentiate the district's adopted reading curriculum.

The fact that Central is highly committed to teacher learning and development came through in all the interviews that were conducted. Veteran and novice teachers, as well as district administrators, spoke of the ways the district supported and encouraged teachers' learning opportunities. Many teachers reported that the professional development they receive currently is of a "much higher quality and much better because there's a focus to it." Another teacher added, "And it's not the latest trend that comes along; it's research based. It's known to be successful."

Additionally, the coordinated effort towards professional development has allowed space and time for teachers to work together. Often times, the district-sponsored professional development brings grade-level teachers together to work on one topic such as differentiated instruction. This allows the teachers to collaborate, share ideas and materials, and find out what is going on at other sites. One teacher remarked, "They always set a pretty decent block of time aside so that we really have some time to work together as a whole group and then work together as site groups. You know, so we can take what we're working on and really use it."

The superintendent views teacher quality and ongoing teacher learning as inextricably linked when faced with the task of raising student achievement. He stated, "When all is said and done, and the doors in the classrooms get shut, it's about the teacher." Thus, the district supports teacher development and learning through a focused staff development program. As one

teacher shared that since it is more of a focused effort, “I feel like the development that goes on is not wasted. It’s not duplicative or anything like that.”

### *Social Capital: Instituting the Coaching Model*

To further the district’s professional development initiative, a coaching model was implemented as a vehicle for instructional change. A central office administrator reported the district’s philosophy of providing classroom teachers with coaches to improve teachers’ skills and their practice. Significantly, the coaching model is not evaluative. The teachers and district personnel reported that it is strictly viewed as a support mechanism for teachers. As such, the coaches are invited into classrooms by teachers rather than mandated to work with teachers as an extension of the administration. The coaching model was instituted in 2003 with one coach. Since then, the model has expanded to include six full-time coaches with each having their own specialty: technology, K-2, 3–6, Gifted and Talented Education, English Language Development, and secondary focus. The coaches meet regularly with each other and the director of curriculum and instruction to talk about the needs of the teachers and students in the district. These meetings contribute to the planning of the district-wide professional development that is offered to the teachers.

Modeling lessons is a large component of the coaching model. The theory behind modeling lessons is that observing others can be a powerful opportunity for learning and enables teachers to get more targeted, individualized support (Neufeld & Roper, 2003; Poglinco et al., 2003). As the teacher watches the coach demonstrate a model lesson, this has the potential to give the observing teacher insights into good teaching practices and strategies. In addition, observing a coach model a lesson sometimes validated the struggles the teacher was having in the classroom. The director of curriculum and instruction stated, “Sometimes it was enough knowing it doesn’t always work for the coaches either. It’s that teaching is hard business. ‘Oh, okay. It doesn’t work for them either.’ They [the coaches] have powerpoints that go off in the middle of something. They have glitches. But they keep persisting.” The principal reported that she found that many new teachers take advantage of the opportunities to work with coaches. In her opinion, the coaching “gets them on the right page faster, instead of having to flounder ... It gives them a safe person to call that can get them better on the right page earlier on, I think, in their career as a teacher.”

The coaches and the director of curriculum and instruction explained that they are in place to support the teachers and their work in the classroom. The focus group teachers expressed that the support shows up as creating teacher-made materials to supplement the district's curricula. As an example, a teacher explained how the materials have made teaching the reading program so much easier. The coaches made picture vocabulary cards to go along with the reading program's stories, for each grade level. The focus group teachers indicated that this has been particularly helpful in their work with English Language Learners. Although this act of providing teachers with teaching materials may seem trivial and not likely to build teacher knowledge or strengthen practice, it appeared to serve as a big source of support for overburdened teachers. Providing teaching materials, therefore, served to strengthen the coaches' roles as one of support rather than evaluation and build their credibility among the teachers.

The coaches explained that they become the conduit to other teachers throughout the district so that "everyone benefits." The director of curriculum and instruction expressed, "There's a lot of great things happening in our own classrooms. And it's getting that message out and then sharing it. That's breaking down the walls. And I think the coaches are helping do that." Through the coaches' sharing of materials and ideas, individual teachers are being praised and recognized for their creative endeavors and hard work. A coach commented, "Within the district, the communication level has really gone up. And ideas or things that they want to share or things that we're sharing with them, we come across a good idea for a particular grade level, we'll send it to all teachers in that grade level. So they all have access to it." The Central website extends the ways in which the coaches share teachers' work and ideas. Creative lesson plans, ideas, and teacher-made materials are placed on the website. Presumably, district-wide access allows for much more cross-school sharing ensues. In addition, the technology coach reported that he receives emails from teachers in other school districts who comment on how useful the teacher ideas are.

The coaching model provides teachers opportunities to see other teachers teach. The coaches organize small groups of teachers to take a tour of four or five classrooms throughout the district. The teachers are provided a substitute and released for the day to visit their colleagues' classrooms. After the tour, the teachers spend the afternoon talking among themselves and the coach about what they saw, their own practice, and ways to incorporate new ideas and strategies. The director of curriculum and instruction explained that the "reason we like taking them in our own district is we want them to see that whether you go to a high socio-economic school or a low, that the

challenges are the same and that there are gifted teachers in your own grade around here that you can email or ask for ideas.” The coaches agreed that the teacher site visits can be a very validating experience. Seeing someone else teach the same way or struggle with the same issues demystifies teaching. This empowers the teachers to go back to their classrooms and say, “that’s something I can do” or “I can modify it this way.”

### *Building Trust*

A key component of the coaching model is trust. To begin, the coaches must be invited in to the classrooms. The director of curriculum and instruction stressed, “If you don’t want them to come in and model a lesson or give you feedback, no big deal.” Once the coaches are invited in, anything that is talked about is strictly confidential. A coach shared her perspective, “The teachers realize that we’re not coming in and evaluating them. We’re not judging ... They invite us in because they know it’s not going to somehow end up on an evaluation someplace.”

However, getting teachers to view the coaches as not part of the administration was difficult at first. A coach reported that the biggest challenge can be battling perceptions that since you are a coach and housed in the district office, now you are “one of them.” He stated, “The trust level I think is the hardest part at the beginning – is building that trust level. I mean depending on the group you’re at, it takes a long time to earn that. And you do it one person at a time.” In order to combat misconceptions of their role, the coaches make it a point to spend generous amounts of time at the school sites talking with teachers, eating lunch with them, and becoming more familiar to them.

Further, the coaches, teachers, and administrators agreed that the success of their coaching model hinged on the fact that the coaches are well-respected teachers who have taught in the district. As a result, the coaches are perceived as peers. One teacher reported, “These teachers have been in the classroom and they know what it’s about.” Another teacher added, “They are people who have been around for a long time and I think that really is the buy-in that a lot of us have right now. They’re people that have been in the district and have been in the classroom.”

The district administrators seem to highly respect and hear what the teachers have to say. In turn, the focus group teachers expressed feeling listened to by the district. A coach shared, “I believe the district really listens to what teachers and students need, and tries to meet those needs in the best

way possible.” This was evidenced at the district-sponsored professional development that was observed. During the workshop, the teachers were encouraged to write comments, questions, and suggestions on index cards and pin them up on the “parking lot.” At the lunch break, the director of curriculum and instruction stopped by to have lunch with and talk to the teachers. Before he left the session, he gathered the questions and comments that dealt with administrative issues and concerns from the parking lot such as how schools are to test and classify their English Language Learners. He returned at the end of the day to address the teachers and to respond to their comments and questions. For those questions he did not have an immediate answer to, he stated who he would contact to discuss the issue and gave an estimated timeframe for how long the teachers would have to wait before they received a response. It was clear that the teachers felt heard. One teacher reported, “What most teachers say about him [the director of curriculum and instruction] is that he also listens to us.” Bryk and Schneider (2002) suggest that feeling heard is an indicator of relational trust. In addition, a study conducted by the National Center for Education Statistics (1997) showed that working conditions such as administrative support and leadership are positively associated with teacher job satisfaction. Although the example described above does not demonstrate the direct effects of trust on teacher learning, it does highlight the need for the development of trust across all levels within a school system when implementing ambitious instructional reform. As district administrators often make policy decisions that affect teachers’ work, trust must be developed up and down the system as well as across the system in order for substantive changes to occur within the classroom.

The relational trust developed in the district can also be attributed to the existence of administrative support from the highest level. The coaches reported that their work is not just backed by the director of curriculum and instruction, but also by the assistant superintendent above him and the superintendent above her. From their lead, support for the coaching model has spread to the site level administrators too. A coach noted, “I know from previous experience that if the site administrator is not going to support it, it’s not going to make any difference how much money you want to throw at it.” As a result, the administrative support allows coaches and teachers to work around the weaknesses they perceive in their selected literacy curriculum. Therefore, the curriculum is not rigidly implemented. Teachers and coaches are given autonomy to do what they feel is in the best interest of their students. Further, as a nod towards uniform support, all of the district administrators went through the same 40 h, AB466 training as the teachers.



This left the teachers feeling well supported and feeling as though the administrators better understood their work.

*Physical Capital: Forging Organizational Changes to Support the District's Goals*

A “changing of the guard” ensued in 2001 when the current superintendent arrived. The superintendent brought in a lot of new leadership and created new management positions. The superintendent described the importance of the new leadership as being, “the most fundamental fact.” It is vital to “have committed leaders who are truly leading. Because if you don’t have the leadership all that other stuff is not-you can’t do it.” For example, the director of curriculum and instruction was originally hired for a newly created position, the director of professional development. This brought on an “intense” focus on professional development in the district that aligned with the needs of the students. A teacher commented on the hiring of the director of professional development, “In my mind, that’s when things started to change. So I would say it came from the top down. It was his [the current superintendent’s] vision. He hired the director who was basically told, ‘It’s your job to improve the academic standing of our schools’ and the director came up with his plan and they supported it.”

In response to the superintendent’s charge, the then director of professional development instituted the coaching model to support teachers’ ongoing learning and development. The director reported that he modeled the coaching initiative off from the BTSA program. He realized the teachers’ need for coaches. “You need to have someone who helps you and will give you feedback. So one of the first things that I did when I came on board was I convinced them to hire a couple of literacy coaches, teachers on special assignment, and now we’ve expanded that over the last couple of years ... But we have teachers on special assignment who do nothing but coach, model lessons, go in and watch lessons, give teachers feedback. It’s not evaluative. It’s peer to peer.” Thus a coaching staff was created in order to support teachers’ development and raise student performance. Teachers with complementary skills and backgrounds were chosen to be a part of the coaching team. Throughout the district, the coaches are known for their ability to work with children. One teacher explained, “These are people who have been in the district forever.” Another teacher added, “They’ve been in our classrooms and they know those grade levels and they focus on them. That’s where they put all their energy. It makes a big difference.”

Other changes were soon to follow. While at one time the district employed a track schedule, the administrators chose to move to a traditional, nine-month schedule. This move aligned more with the district's push towards a district-wide professional development program. A teacher commented, "Now that all of our schools are traditional, we're allowed to have everybody on the same schedule so that we can do the training at the same time where before you'd have people off track, on track, so you had to do it a few different times. And it's nice when we get together with fourth grade teachers, that you get to talk to each other, the collaboration between schools. Because even though we may have it at the same site, you don't necessarily get it at all the schools and find out what's going on."

Further, the district has implemented a district-wide meeting schedule. This organizational change has provided more opportunities for ongoing staff development. Thus, at Central, every Monday afternoon is reserved for a school site staff meeting or a grade-level meeting. The director of curriculum and instruction reported, "The idea was to build a common structure." This common structure has allowed teachers to get together as a group and given them a consistent venue to talk, share ideas, and problem-solve issues. A coach related, "It's guaranteed every other week that you're going to have department or grade level time available where before it was like two half days a year at the elementary and that was it. You know, unless you wanted to meet after school on your own time." In addition, the district-wide meeting schedule provides coaches with opportunities to present short in-service trainings during those meetings. This, in turn, increases the frequency and consistency of learning opportunities offered to teachers.

The superintendent highlighted the district's commitment to prioritizing literacy development for all students, particularly through resources and funding. Initially when they realized that the cost of their textbook selection was more than they anticipated, the district's literacy priority fueled the administrators' decision to find a creative solution to the funding problem, which they ultimately did. The central office administrators reorganized the budget to deliver the needed resources that would enact their vision for teaching and learning.

Teachers are provided with supplies, textbooks, and materials that are needed to make the district's literacy priority an attainable goal. One teacher remarked, "It definitely has to take an investment from the district as far as monetarily goes. I mean the supplies and stuff that have been offered to us and the workbooks and stuff that have been run off and the materials that we've been getting, I know had to cost a lot of money. And so we've really

been fortunate with this program [reading curriculum] that they've really tried to give us everything that they thought was essential to making the program work."

Funding has also been invested in paying for substitutes to teach classes while the classroom teachers attend district-sponsored professional development. The superintendent shared that the district pulls together federal funding, state categorical funding, and district general funds in order to provide teachers with high-quality staff development. From his perspective, "The bottom line is, we cannot produce results if we don't spend what we have to spend in staff development. So when she [assistant superintendent of educational services division] says she has to have it, or she needs it, we have to provide it." Thus, financial resources were never viewed as barriers to the district's goals. Instead, the need for funding created opportunities for the district to realign its priorities and to illustrate the district's commitment to building the capacity of its teachers.

## DISCUSSION

This case study illustrates the influential role that district leaders can play in building the capacity of teachers. District administrators engaged in a series of important leadership roles. First, the district officers created a coordinated vision for improving teaching and learning in the district. This vision led to the development of a creative approach to ongoing teacher professional development, the coaching model. Finally, leadership in the district office made changes in their organizational structure that supported the district's coherent vision for instructional reform and reinforced it with the necessary funding. As one teacher put it, "They put their money where their mouth is." This coherent focus enabled the district to build the capacity of its teachers and improve student learning.

The prioritizing of the district's goals and the administrators' attention towards creating a district-wide teacher professional development model can be viewed as human capital development (Spillane & Thompson, 1997). The district's shifts in priorities correspond to features that have been found in other research to be associated with changes in teacher knowledge and practice. For examples, collective participation in teacher professional development has been found to be more effective than individualistic participation (Cohen & Hill, 2000; Garet, Porter, Desimone, Birman, & Yoon, 2001). The purposive focus on increasing teachers' knowledge and skills as well as cultivating their dispositions towards instructional reform

was essential for enabling changes in instructional practice and furthering teacher learning and development. The teachers' positive views of district priorities and efforts to support high-quality teaching and learning suggest that the district administrators have advanced a professional culture (McLaughlin & Talbert, 2003). This culture positively affects teachers' motivation and willingness to engage in the reform.

Central's use of the coaching model provides evidence of the initial steps taken by the coaches to foster social capital among its teachers. Coaching, as a mechanism for professional development of teachers, is a form of inquiry-based learning and centers on collaboration between teachers (Poglinco et al., 2003). The practice involves ongoing classroom modeling, observations of teaching practice, and individualized feedback that encourages the modification of instruction and the support of new teaching methods. The district developed social capital by creating a shared culture of instructional improvement and providing teachers with a network of social support from colleagues to whom they could seek for help. This, in turn, was key to the development of relational trust within the district.

This study supports Bryk and Schneider's (2002) conclusion that trust is a necessary resource for school improvement. Relational trust is the synchrony of obligations and expectations among those within a community and the level of trust catalyzes organizational and instructional improvements. Within an organization, trust relations allow for more effective decision-making, enhanced social support for innovation, and an expanded commitment to the organization. The effects of relational trust were evident in Central Unified School District. The teachers' descriptions of the district administrators often emphasized the administrators' personal integrity and competence. The teachers revealed statements about how they felt heard by district leaders regarding organizational and pedagogical issues and concerns. Because the district administrators responded to the teachers' concerns, they felt supported by the administration. As a result, this went a long way towards garnering teacher buy-in for the district's vision for instructional reform.

The process of human and social capital development among teachers is intimately related to the physical capital available within the district. Within Central, not only were necessary funds made available to ensure the success of the district's priorities, attention was paid towards providing teachers with time to learn new instructional methods. For example, teachers were given release time and substitute teachers were brought in to allow teachers to attend professional development sessions or to observe other teachers' classrooms. In addition, the instructional coaches provided teachers with materials to ensure the successful implementation of new teaching strategies.

Although physical capital is not sufficient to bring about ambitious instructional reform, along with human and social capital, it is a vital component in building local capacity.

The greatest limitation of this study is the small sample of teachers that were interviewed. The research team spoke with a greater number of district personnel than school personnel. In addition, while we encountered an overwhelmingly positive response from teachers and administrators about the coaching model, we did not speak to every teacher and administrator. Thus, we do not have a definitive sense of how much the coaching model has permeated teachers' classrooms since it is predicated on the notion that the coaches must be invited in. As a result, we do not know how struggling teachers who do not seek coaching help receive professional development.

Central Unified School District's strategic focus on building the capacity of its teacher workforce laid the groundwork for the innovative approaches that were employed in efforts to raise student achievement for all students. This case study adds to what we know about the strategic and significant roles that central office administrators and staff can play in building the capacity of the teacher workforce. In addition, this study emphasizes the role of professional development in producing both human capital and social capital within school districts. This study contributes to research on school districts by highlighting the particular mechanisms (i.e., district leadership, coaching model) that have the potential to influence teaching and learning in schools. While I am unable to attribute the particular gains the district has made recently on their API scores to the district's instructional reform, it must be taken into account. Thus, there are potential lessons to be learned from Central.

That teachers do value and seek the guidance and support of instructional colleagues, however, is the focus of this chapter. As such, building the capacity within a school district by investing in the development of its teacher workforce is a noteworthy goal. The learning opportunities that are offered to teachers are symbolic of the learning opportunities that are offered to students. Without increasing the knowledge and skills of the instructional leaders in the classroom, one cannot expect similar gains in the knowledge and skills of students.

## NOTES

1. This is a pseudonym for the school district. The central valley is predominantly an agricultural area but does have some urban centers.

2. School statistics were derived from <http://data1.cde.ca.gov/dataquest> and [www.greatschools.net](http://www.greatschools.net) on December 5, 2005.

3. California's state accountability index measures student achievement and school growth in reading, math, history-social science, and science.

## REFERENCES

- Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In: L. Darling-Hammond & G. Sykes (Eds), *Teaching as the learning profession: Handbook of policy and practice*. San Francisco, CA: Jossey-Bass Publishers.
- Bidwell, C. E., & Yasumoto, J. Y. (1999). The collegial focus: Teaching fields, collegial relationships, and instructional practice in American high schools. *Sociology of Education*, 72(4), 234–256.
- Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33(8), 3–15.
- Borko, H., & Putnam, R. T. (1995). Expanding a teacher's knowledge base: A cognitive psychological perspective on professional development. In: T. R. Guskey & M. Huberman (Eds), *Professional development in education: New paradigms and practices* (pp. 35–65). New York, NY: Teachers College Press.
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York: Russell Sage Foundation.
- Cohen, D. K., & Hill, H. C. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, 102(2), 294–343.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–S120.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, 24(2), 81–112.
- Elmore, R. F., & Burney, D. (1997). *Investing in teacher learning: Staff development and instructional improvement in community school district #2, New York City*. National Commission on Teaching & America's Future, Washington, DC.
- Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: The case of computer technology in schools. *Sociology of Education*, 77(2), 148–171.
- Fullan, M. (1985). Change processes and strategies at the local level. *The Elementary School Journal*, 85(3), 391–421.
- 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), 915–945.
- Hawley, W. D., & Valli, L. (1999). The essentials of effective professional development. In: L. Darling-Hammond & G. Sykes (Eds), *Teaching as the learning profession: Handbook of policy and practice* (pp. 127–150). San Francisco, CA: Jossey-Bass Publishers.
- Little, J. W. (1993). Teachers' professional development in a climate of educational reform. *Educational Evaluation and Policy Analysis*, 15(2), 129–151.

- Mac Iver, M. A., & Farley, E. (2003). *Bringing the district back in: The role of the central office in improving instruction and student achievement*. Report no. 65. Center for Research on the Education of Students Placed At Risk, Baltimore, MD.
- Marsh, J. A., Kerr, K. A., Ikemoto, G. S., Darilek, H., Suttorp, M., Zimmer, R. W., & Barney, H. (2005). *The role of districts in fostering instructional improvement: Lessons from three urban districts partnered with the institute for learning*. Santa Monica, CA: RAND.
- Massell, D. (2000). *The district role in building capacity: Four strategies*. Policy Brief No. RB-32-September 2000. Consortium for Policy Research in Education, Philadelphia, PA.
- McLaughlin, M., & Talbert, J. (2003). *Reforming districts: How districts support school reform*. Seattle, WA: Center for the Study of Teaching and Policy.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- National Center for Education Statistics. (1997). *Job satisfaction among America's teachers: Effects of workplace conditions, background characteristics, and teacher compensation*. U.S. Department of Education, Office of Educational Research and Improvement, Washington, DC.
- Neufeld, B., & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity*. Washington, DC: The Aspen Institute Program on Education.
- Poglinco, S. M., Bach, A. J., Hovde, K., Rosenblum, S., Saunders, M., & Supovitz, J. A. (2003). *The heart of the matter: The coaching model in America's choice schools*. Philadelphia, PA: Consortium for Policy Research in Education.
- Richardson, V. (2003). The dilemmas of professional development. *Phi Delta Kappan*, 84(5), 401–406.
- Smylie, M. A., Allensworth, E., Greenberg, R. C., Harris, R., & Luppescu, S. (2001). *Teacher professional development in Chicago: Supporting effective practice*. Chicago: Consortium on Chicago School Research.
- Spillane, J. P. (1996). School districts matter: Local educational authorities and state instructional policy. *Educational Policy*, 10(1), 63–87.
- Spillane, J. P., & Thompson, C. L. (1997). Reconstructing conceptions of local capacity: The local educational agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis*, 19(2), 185–203.
- Sykes, G. (1999). Teacher and student learning: Strengthening their connection. In: L. Darling-Hammond & G. Sykes (Eds), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151–179). San Francisco, CA: Jossey-Bass Publishers.
- Togneri, W., & Anderson, S. E. (2003). *Beyond islands of excellence: What districts can do to improve instruction and achievement in all schools – A leadership brief*. Washington, DC: Learning First Alliance.
- Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professional development. *Review of Research in Education*, 24, 173–209.
- Woody, B., Bae, S., Park, S., & Russell, J. (2006). *Snapshots of reform: District efforts to raise student achievement across diverse communities in California*. Working Paper no. 06-2. Policy Analysis for California Education, Berkeley.

## CHAPTER 8

# EXIT EXAMS AND ORGANIZATIONAL CHANGE IN A VOCATIONAL HIGH SCHOOL

Michèle Schmidt, Marisa Castellano, Athena Tapales,  
Sam Stringfield and James R. Stone, III

### ABSTRACT

*This chapter examines one vocational high school's response to a state exit exam. Many states now require high school students to pass an exit exam before graduating, a key element of standards-based accountability reforms. Little is known about how educators and students inside vocational schools respond to these exams which typically emphasize literacy and academic skills. We examine how one such school attempted to respond to demands linked to the exit exam and the state's labeling the school as underperforming. While teachers reported support for state intervention and placing stronger demands on the school, one remedy involved becoming more selective in terms of new students admitted. As a result, tensions arose between academic subject and vocational teachers. Deep frustrations were voiced by several teachers and students about whether preparation was sufficient to ensure a reasonable pass rate. We employ a critical public policy framework to illuminate how this policy shock spurred positive action while penalizing students for years of insufficient preparation in the public schools.*

---

**Strong States, Weak Schools: The Benefits and Dilemmas of Centralized Accountability**  
**Research in Sociology of Education, Volume 16, 209–234**  
**Copyright © 2008 by Emerald Group Publishing Limited**  
**All rights of reproduction in any form reserved**  
**ISSN: 1479-3539/doi:10.1016/S1479-3539(08)16008-6**



## INTRODUCTION

The implications of high stakes exit exams (EE) on graduation rates are increasingly debated across many nations. More than half of the states require a student to pass a high school exit exam (Mendez, 2004; Peterson, 2005). Arguments against exit exams include concerns that they cause students to drop out, widen the gap between middle-income and low-income students, and raise the question of rigor as some state exams are simplified to ensure that larger shares of students pass. Students of color, along with those with disabilities, typically fail at higher rates than white or middle-class students (Peterson, 2005; Mendez, 2004).

Advocates of such high stakes exams argue that they place productive pressure in educators and students alike to reach more rigorous standards. The benefits include individual attention and extra resources devoted to students. In addition, the exams have exposed the achievement gap between low-income students and other students. Exit exams allegedly prompt stronger training of teachers and administrators and for the development of curriculum and instruction and instructional guides that align with the test (Castruita, 2006).

Case studies on the implementation of exit exams conducted in Virginia and Maryland reveal that the exams can promote discussion of student performance and cooperation among teachers. Yet these exams may diminish teachers' emphasis on higher-level skills and teaching to the test (Center on Education Policy, 2005). Advocates in general highlight the fact that students have numerous opportunities to retake the test, sometimes up to 11 times. Students may be offered remedial and tutorial assistance to meet test requirements (Peterson, 2005). The argument in favor of training for test-taking skills seems to be that it provides time and resources for low-achieving students who normally are overlooked in the system, often in danger of dropping out (Greene & Winters, 2004). But only 11 states allocate funds for tutoring and additional instruction for students who fail (Center on Education Policy, 2005; Peterson, 2005).

With the continued failure of many students, legislators continue to revisit the requirements of exit exams by recommending delaying their implementation or introducing alternative measures (FairTest Examiner, 2007; Peterson, 2005). For example, Washington state has long delayed implementation as a result of widespread failure on the exit exams. In the meantime, end-of-course exams in algebra, geometry, and biology will replace graduation exams (FairTest Examiner, 2007). In other states, such as Arizona and California, withholding diplomas was postponed for

three years. In Arizona, 64% of students in 2004 failed the state math exit exam and 41% failed the English exam (Peterson, 2005). California is currently investigating alternatives to the exit exam for students with disabilities, since only 48% of disabled students passed the test in 2004, and one-fifth of seniors in the regular stream were denied diplomas (FairTest Examiner, 2007).

## CASE STUDY SCHOOL

While policy debates on this topic tend to focus on comprehensive high schools, we focus on how one state's exit exam is understood and responded to within a vocational high school (VHS). This chapter discusses the impact of the EE during the same year (2003) that the school was deemed underperforming by the state. The institution, VHS, enrolled 1,440 students in 2003.<sup>1</sup> The student population was comprised of 13% White students; 58% Hispanic; 28% African-American, and 1% Asian.<sup>2</sup> Sixty-eight percent of students were eligible for a free or reduced lunch with 45% ESL students, 10% Limited English Proficient, and 30% special education students. The attendance rate in the school was the lowest in the district (Table 1).

At VHS the failure rate on the English component of the EE increased from 89% to 95% between 1998 and 2000. In 2001, the failure rate in math rose to 92% by 2002. Throughout the five-year period from 1998 to 2002, the school's overall failure rate was higher than district and state averages (School Panel Review Report, 2002).

**Table 1.** Descriptive Data from Vocational High School, Its District, and Its State.

	Vocational High School	District	State
Total number of students	1,400	26,000	1,000,000
% Latino	58	45	11
% African-American	28	29	9
% White	13	23	76
% All others	2	3	5
% Free/reduced price lunch	68	66	25
% Limited English Proficient	10	11	5
% Special education	30	21	16

*Note:* Data are from 2000–2001 school year. Data have been rounded in order to ensure anonymity.

In 2005, the school once again was in jeopardy of being declared underperforming. Yet, by 2007, only 35% of students failed the math component of the EE.<sup>3</sup> This dramatic shift illustrates the difficulty in assessing whether the structural and individual changes detailed in our analysis were truly effective. It appears that the state's EE had become easier to pass (Table 2).

This policy complication is not unique to the state in which VHS is located. For example, in 2002–2003 school year New York discounted the results of a new math test for junior and seniors after the passing rate fell much lower than the one of the previous year. Local officials received permission to give diplomas to seniors who failed the exam but passed their math courses (CNN, 2003). In Massachusetts, where diplomas were withheld for the first time, some students walked out of class and refused to take the test, often with support from parents. Most states gave students another route to obtain a diploma. Florida, for example, allowed seniors who failed the state test to substitute another one, such as the SAT.

Still, little is known about how vocational schools attempt to respond to high stakes testing, which focuses largely on students' acquisition of academic skills (Austin & Mahlman, 2002; McGinley, 2002). This chapter delves into how VHS, which attempted to balance both an academic and vocational curricula, responded organizationally and individually to its state exit exam. Our research questions are: What is the impact of an exit exam on the deep organizational structure and curriculum of a vocational high school? How do teachers and students respond to the curricular demands of the exit exam.

**Table 2.** High School Exam Scores for Vocational High School, District, and State, 2001–2007.

	Percent of Students Failing the High School Exam in						
	2001	2002	2003	2004	2005	2006	2007
<i>High School Exam – Math</i>							
Vocational High School	89	92	63	46	58	43	35
District wide	65	64	53	39	46	36	31
Statewide	25	25	21	15	15	12	9
<i>High School Exam – English</i>							
Vocational High School	82	76	44	40	42	23	25
District wide	50	64	34	28	30	21	20
Statewide	19	14	11	11	10	7	6

## THEORETICAL TOOLS

We draw from a critical policy framework which emphasizes how power relations inside a school must be examined to understand whether and if so, how the organization responds to a perceived policy shock (Ball, 1994; Scheurich, 1994). This enables us to employ theoretical tools for unpacking how both educators and students within a VHS interpreted a given policy landscape and its policy shocks (Ball, 1994; Taylor, Rizvi, Lingard, & Henry, 1997). Rather than seeing exit exams as technical devices to be adopted, we examine the social forces at play which shape interpretation of and responses to policy (Cibulka & Derlin, 1998; MacPherson, 1996; Datnow, Hubbard, & Mehan, 2002). Such a lens provides a critical perspective, allowing us to address the more enduring and deeper questions of whether schools are making meaningful changes to address (or rebuff) high stakes testing. Our analytic frame also prompted the question: Who determines these changes? By using a critical policy lens, the goal becomes not simply to raise test scores, but rather, the attainment of a schooling process that is in the best interest of every student (Capper, 1993; Sirotnick & Oakes, 1986).

Advocates of high stakes accountability may ignore the social context of vocational schools, as they expect universal responses to a standard policy remedy. Critics claim that high stakes tests can undercut student motivation, result in teaching to the test, and higher drop out rates. Analyses by post-modern or post-structural scholars (Ball, 1994; Foucault, 1980) maintain that policy is actually multi-dimensional in nature, as well as value-laden and contextual. Policies interact and sometimes conflict with each other. Furthermore, they are not always rational and often yield negative, unintended consequences. Ball (1994) and Foucault (1980) stress that policy as discourse becomes a power exercise as to whose meaning is legitimated, whose voices are heard, and whose values are recognized or authoritatively allocated.

## METHODOLOGY

This chapter stems from a larger 5-year longitudinal study of the impact of comprehensive school reform models on vocational middle and high schools (Castellano, Stringfield, & Stone, 2002).<sup>4</sup> The larger study used a mixed method design and involved three cohorts of students in grades 7–11 in 10 schools at three sites over a 4-year period between 1999 and 2003. The sites were compared with matched control sites not implementing the reforms and with three replication sites which involved similar reforms (Castellano et al., 2002).

VHS was selected for the present analysis from the larger study sample. It was part of a cluster that included a comprehensive high school, a high school divided into academies, a rapidly growing structural innovation (Kemple & Snipes, 2000), and a VHS that balanced academic and technical studies (Lynch, 2002).<sup>5</sup>

Some characteristics unique to our study school are relevant to this chapter. For example, curriculum integration was a significant focus in the ninth grade. All students were expected to select a shop by the end of the first 10 weeks of their freshman year as mentioned above.<sup>6</sup> Since one-third of the incoming ninth graders' reading and math skills were far below grade level, a prep school that integrated academic subjects with vocational applications was designed. In addition, a computer-based remediation system helped raise students' reading skills from freshman year proficiency levels to grade level in reading and mathematics. For example, in 1999–2000, 135 of the 160 incoming freshmen that tested below grade level were able to complete the ninth grade curriculum by the end of their freshman year (Castellano et al., 2002). In addition, just prior to our visit, the school had instituted a computer networking shop with the opportunity for A+ certification for all graduates. Students in this shop learned how to build and maintain networked systems, applying that knowledge by maintaining the entire system in the school (Castellano et al., 2002).

VHS had a recent history of gang violence, fearful teachers, and apathetic counselors. Prior to our visit, steps had already been taken by past administrators to reduce gang violence, develop partnerships with community businesses, implement a High Schools That Work network within the school, and offer Advanced Placement courses. Issues within the school that remained problematic, however, were the physical plant, staff capacity for change, student behavior, and academic rigor (Castellano et al., 2002).

One of the advantages of longitudinal studies is the ability to follow-up on emerging themes, one of which serves as a point of discussion in this chapter. While the goal of this chapter is not to explicitly comment on whole-school reforms as was the intent of the larger study, rather, we decided to follow-up on an unexpected and emerging theme that we soon realized was a prevailing theme across the U.S. – the influx of failing students who are unable to attain their graduation diplomas. Our interest in this phenomenon was piqued when we began to note lagging implementation of this reform, coupled with the fact that the school was the only one in the district to be deemed underperforming and required state intervention due to the high numbers of seniors failing to receive their graduation diploma.

### *Sample*

The sample of individual study participants were drawn from the student, teacher, district, and administrator population at the school. The student composition of our study school was suburban and poor and included a large number of minorities. The sample was comprised of 8 students; 28 teachers (12 academic, 16 vocational); 2 district administrators (i.e., superintendent, assistant superintendent); 1 principal; 1 secretary; 3 guidance counselors; and 8 coordinators (e.g., bilingual, special education, reading, and vocational programs). For this present chapter, we include responses from the district superintendent, the principal, two guidance counselors, three academic teachers, five vocational teachers, three coordinators (e.g., special education, read/write, and bilingual), and four students, all of whom spoke directly to the topic of this chapter concerning the impact of EE within the school.<sup>7</sup>

## **CASE STUDY**

Using a case study approach, we “investigate a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident” (Yin, 1994, p. 13). Of all the schools in the larger study, our study school was the only vocational school and the only school in which the state intervened due to low EE scores and high numbers of failing graduate students. While a key criticism of case study methodology is its dependence on a single case to render findings not generalizable across other settings, Yin (1993) and Hamel, Dufour, and Fortin (1993) maintain that the importance of case studies in describing, understanding, and explaining should not be underestimated on a local or global level. We believe that while this chapter tells only one story of one school’s experience with the exit exam, it contributes to the growing body of literature on the impact of exit exams particularly for educators and students in a vocational program.

### *Data Collection and Analysis*

Qualitative data were collected through interviews with participants, conducted during a two-day visit to the school. Two person teams conducted the interviews with each participant. Through semi-structured interviews, the study team asked the participants about their perceptions of the accountability

policy, their response to probationary status, and changes made after the school was placed on probation. Our descriptive analysis highlights themes and issues that arose along with participants' views of accountability and their response to the policy. Although these data do not warrant generalizations to other populations or schools, they are critical to our understanding of both the policy's theory of action and the theories related to intended and unintended consequences of a policy within one school.

Interviews with individuals were taped and transcribed. The first stage of the coding and data analysis involved reading the transcripts and marking the instances of themes of interest to this chapter. Using coding methods according to [Strauss and Corbin \(1990\)](#), we began with open coding, by identifying data relevant to the EE policy, which were then grouped into categories using axial coding ([Strauss & Corbin, 1990](#)). These codes were imported into a qualitative software program (HyperResearch, version 2.6). As an example, we coded references to how VHS staffs were preparing students to take the EE. These instances of various kinds of test preparation were grouped by category, such as tutoring students, or engaging in item analysis of test questions. These categories were then axially coded into CTE classes and academic classes, in order to determine the extent that EE preparation had permeated the classroom.

Conceptual coding proceeded in a similar manner. However, the categories reflected the theoretical ideas found in our conceptual framework. Implementation analysis grounded in post-structural theory enabled us to identify the transformative processes attached to stakeholders' experiences. We were able to highlight difficult and key questions and reflections that the various stakeholders had about their experience, including where teachers and administrators recognized that the press for high stakes testing involved power struggles over what forms of learning should be awarded high priority. Changes within school structures and in the beliefs of stakeholders are often influenced, and in some cases undermined, by political and bureaucratic interference or institutional priorities and requirements that obfuscate significant changes in the implementation of testing policy ([Earl & LeMaheiu, 1997](#); [Firestone, 1998](#); [Hargreaves, Earl, & Schmidt, 2002](#)).

## ONE SCHOOL'S STORY

The school district and VHS are made up of a suburban population with a large number of minority residents. The school began as a traditional VHS in the 1920s. Over the years, it served the needs of the community as a trade

school where students received both academic and vocational training and it remains the only VHS in the city, along with three other high schools dedicated to college preparatory curriculum. Gradually, however, VHS shifted to serving a student population comprised of ever more “special needs” students. The city, historically a manufacturing center, experienced white flight in the 1980s. The poverty rate of the city grew to over 40%, which was reflected in the number of students who participated in the free and reduced-price lunch program. The percentage of special education students at VHS was one of the highest in the state at the time of our study.

The state in which our study was conducted had a long history of local control of schools. Graduation requirements and course content were historically left to the discretion of local school boards and districts. The State department education traditionally certified teachers, provided professional development, and coordinated state standardized testing. In the 1990s, as part of its educational reform movement, the state instituted a new performance-based assessment system. Compliance was monitored by the state.

We began to see dramatic changes to school structures (e.g., timetable, schedule, curricula, programming, and mission statement). Test results were used by the state to publicize a failing ‘report card’, as it was declared underperforming. Typically, school districts use such tests to decide which students are qualified to be promoted from grade to grade and who is qualified to graduate from high school.

During the study year, VHS was declared to be underperforming by the state due to its continued poor performance on the EE. In addition to dealing with this in-crisis declaration, the staff of VHS was also working to address other issues, which included district neglect of the physical plant.

Before the state took note of the declining numbers of graduating students and test scores, the EE had already been administered for a few years. Exit exams were administered annually beginning in elementary schools and continuing through high school. The goal of the test is to improve teaching and learning by encouraging teachers to identify strengths and weaknesses of students, as well as motivate parents to monitor their children’s progress. Just prior to our visit, the EE had become high stakes, with tenth graders required to pass the Math and English sections of the EE in order to graduate from high school, with additional opportunities to pass before the end of their senior year. Tests scores are reported to individual students, schools, and districts according to performance levels defined by the Board of Education.

However, the year the EE became a mandatory graduation requirement, it became evident to school administrators and the state that the school’s test



results were poor, and student and teacher attitudes toward the tests apathetic. When the school was designated underperforming in 2003, this designation caused the faculty and staff to adjust the curriculum and change the class schedule to accommodate more time and resources in test preparation. This resulted in more time being spent in academic courses and less in vocational courses. The already distant relationship between academic and vocational teachers became more strained given these adjustments. Both parties acknowledged the importance of preparing students for the EE.

## A FOCUS ON ACADEMIC ACHIEVEMENT

At the time of our visit, teachers were clearly aware of the sanctions in place if the test scores did not rise. For example, one CTE teacher, who was concerned about keeping his job as more and more emphasis was placed on academic curriculum, commented:

I'm a pragmatist. I know that if I don't help my students raise their standards, I'm not going to have a job. Because vocational education is not going to be what it once was. It's becoming more comprehensive. And if this building is an under performing school, and it doesn't improve, the state department comes in, as I understand it, and they literally dismiss every staff member and hire a new staff. And I don't want to do that.

Teachers' daily lives revolved around the goal of raising student achievement levels. Teachers and the principal were anxious, but also hopeful about achieving their goal of increasing literacy levels, which would ultimately raise test scores. Little was spoken about 'raising scores' per se, rather, the objective became 'increasing literacy levels'. The principal expressed it thusly:

Well you know the underlying goal that drives both math and English at this particular point for our school, and I expect that that'll change as the school evolves a bit. But the underlying goal is building reading comprehension, reading literacy. That is probably the single most important factor in our school to boost achievement.

The staff rallied together and created a culture of hope and pragmatic realism. One teacher commented on the staff's cohesiveness:

I think we know we have a really hardworking, dedicated principal. And we feel as though he is doing what he can to make this a better school and the people who are concerned, the good teachers in the school, which is a large percentage. My feeling is that of all the schools in [this city], I think we have more people here who are genuinely altruistic. We have a core of these teachers here. A really strong core. And if you can tap into that, and make a difference...that would be the best thing to create a change in culture in the building. And I think that's happening.

## **RESTRUCTURING TEACHING TIME**

We learned that a number of scheduling and timetable changes had already occurred in the past year, with the hope of carving out more time to prepare students in taking the state test. Specifically, the school timetable changed from a 5 period day to a 6 period day with the sixth additional period intended to provide EE-assisted courses. In addition, the schedule changed from a 10 week cycle to a 5 week cycle, as students had 5 weeks in shops and 5 weeks in academic classes, giving the ninth graders a year long split schedule that was divided between academic and vocational classes. In addition to these timetable changes, a new curriculum was introduced into the English and math EE-assisted classes. As one teacher comments:

So now they're getting double the math, double the English, besides the helping courses being put in there: additional math and additional read/write or English. So they're kind of being bombarded in ninth and tenth grade, which makes sense. Get them really prepared with the fundamentals before they take the test.

Students had two daily blocks that concentrated specifically on a shop-related English and math courses. As a result, students lost two shop periods to accommodate these test-assisted and shop-related academic classes. Not surprising, many shop teachers expressed some concern at the loss of a vocational focus, as these changes had profound structural and conceptual implications for the vocational side of the school.

One shop teacher explained that students would “be spending twice as much time doing the textbook work related to the shop. Cosmetology has a book; culinary has a book, so you're getting double the bookwork now in the shops.” Another shop teacher remarked, “Because of the exit exam and the being declared underperforming, they concentrated so much on the math and English and the academic part, that the vocational is being hurt ... So I'm actually losing time because I have to teach my related [academic] during shop.” Conversely, one shop teacher displayed less concern about the loss of shop time as teachers were expected to incorporate the reading and writing curriculum in their classes by saying that he was already using about half of his class time to focus on reading and writing:

I really feel that probably 50% of what I teach is really [text] related and not so much hands-on, you know based on the competencies ... [So] we didn't really lose two period. They always [had reading/writing for] one period anyway ... So actually what I've lost in time with the student, is probably 45 minutes a day.

All students were required to take a Read/Write course in ninth and tenth grade “to improve their reading and writing skills.” Another teacher

commented: “We found that [students] had trouble reading and comprehending the questions on the EE. So if you can’t understand the question, you can’t answer it correctly.” Yet another teacher said: “We implemented a couple new programs at the ninth grade level and tenth grade level to get students up to where they would be, in time, to possibly pass the EE in tenth or eleventh grade by extending math for a whole year. Forty weeks of math; 40 weeks of English.” In the past, these courses were 20 weeks in length.

A teacher referred to the 5/5 cycle of academics and vocational classes by saying that this length of time was suitable for both teachers and students since it gave each ‘a break’ from each other as well as resulting in ‘shorter marking periods’.

Whether the kids liked you or not, you needed a break. So to me, I didn’t like that part of it, but you know we saw each other for the whole year. So I like the five weeks on and five weeks off because it’s been a cut off at five weeks and then we’ll have new kids for the next five weeks. So I like that part of it, shorter marking periods.

The curriculum changes related to the introduction of EE-assisted courses impacted both academic and shop teachers alike. One shop teacher said, “I never thought it would impact me directly, and it did because now we have the [writing program] going on in our building. Everyday my students write.” An academic teacher commented that he had to ‘rethink and tighten his planning but also sees the benefit of covering academic material in a shorter time cycle’ (that is, in five weeks) since students are noticeably less bored than with a longer cycle (that is, ten weeks). Furthermore, this tightened cycle kept students focused and ‘on-target’:

Personally, it’s caused me to rethink and tighten up my planning. But I think it’s a better use of a shorter amount of time. I think that the students are less likely to ... be bored in one program or the other. And I think also that the fact that they’re taking, in many cases, two academic classes along with shop, is keeping them on-target more. So I think it’s beneficial, both for the academic side and for the vocational side.

## **CURRICULAR TENSIONS, BUT APPARENT SUCCESS**

This same teacher admitted that he was having difficulty covering the curriculum in his subject area: “I feel I’m struggling to figure out how much I can cover in five weeks, as opposed to ten weeks and trying to block it off.” However, a different teacher believed that the 5/5 cycle promoted better instruction: “I think it’s better.” Teachers in general believed the cycle was better for students. A teacher commented: “[It’s a] good thing for kids. I think

it's better for the kids because of the attention span of the kids." Additionally, one teacher felt that since students were not away from the academic or shop for too long there was less chance of them forgetting what they had learned:

Because you're not away from your academics or your shop for that long a period of time ... [In a longer 10/10 cycle] they've forgotten or they need to kind of brush up. [They might say,] "Oh I forgot how to do this equation." Or I didn't quite remember the characters in this particular play. But if you're not away for such a long period of time, you may have to just do a day's review, instead of almost a weeks review.

Overall then, teachers felt that this new cycle motivated students, leaving them less bored and frustrated. Furthermore, this same teacher said that students who were struggling with academics had a bit of a reprieve from their struggles when they were able to switch to their shop after 5 weeks:

Students that are with an academic that they may be struggling with, know that they only have to put up with that for ten days and then they switch ... And they have a little bit of a reprieve. They can still go back and see their teacher after school, at night, or whatever, work on weekends. But it gives them a little break.

It also appeared that discipline problems decreased as a result of the new cycle. One teacher noted:

So we find that there are more pluses, even for the faculty. They get new faces every two weeks. Cause I've been there. Cause when I was a teacher it was nice to get the new faces. Yeah and to be honest with you, the discipline problems go away for two weeks. You know you get the new ones but at least you have a two week reprieve from that particular student or so on in that class.

Teachers reported having greater control over student behavior. One teacher said:

I think things have tightened up around here, as far as the control of the building. [It] seems to be much better. In other words, teachers are out in the halls. We have hall duty this year. Teachers have hall duty. And which is a real good thing. We haven't had it in the past.

We also learned that there was tutoring available for students after school and on Saturdays:

There are exit exam tutoring programs that are currently being offered and additional programs, we were told, are going to start possibly after Thanksgiving because of grant money coming from the department of Education earmarked for this school, for that one reason: to get us to become a performing school.

The funding that this teacher mentioned was intended "to pay teachers to remain after school to tutor students in exit exam related areas." At the time

of our interviews, success of the tutoring program was minimal, primarily because students rarely remained after school.

## **ERODING STUDENTS' WORK EXPERIENCE**

There were however some negative consequences. For example, the addition of extra EE-assisted classes and a shift to the re-alignment of blocks in the timetable played havoc with students' opportunities to take electives and co-op opportunities that relate directly to their vocational experience. One teacher said:

At the upper levels, student who hadn't passed the EE were forced basically to take ... a second math, and to go into this Read/Write program. So they were told you couldn't take standard electives, art, music, computer. You're going to be steered into taking a math or a Read/Write course because you need this to pass the EE. We were sorry you can't take art, and we have nothing against the arts and music, but you can take those but they're not going to let you out of this school if you don't pass the exit exam.

Another teacher commented that, "The master schedule has been affected. The ripple effect of that master schedule has posed difficulty in placing students on supplementary programs such as any of the School To Work stuff, any of the co-op stuff. It's hard to place kids [in co-op programs]." We learned that other courses designed for the honors students (who make up only a small percentage of the total student population) were almost eliminated, but at the behest of the principal and supporters, the timetable was worked around the honor program:

So we scheduled around them. We didn't say: "Well we can't squeeze in this honors history course with 8 kids. So we're going to eliminate the program and put them in CP. Our principal would not tolerate that. With the true honors kids, they're scoring high. These are potential four year college students, we're gonna implement the honor program and work around the glitches, and we did that. So we worked [it] out.

## **WORKING AGAINST PARALLEL REFORMS**

In addition, we found that the schedule and programming changes have compromised whole-school reform efforts. For example, small learning communities, is a component of the High Schools That Work reform, are now difficult to implement: "With this new schedule, it's impossible [to make the learning communities happen.]" Another teacher explains why the school has decided to double up the amount of math and English students

are now receiving: “It’s actually because these kids only had a half a year each of algebra and geometry and they didn’t get it all. According to the teachers we spoke with, the reality of the situation for these students was that they were trying to catch up with reading and writing that other students at other schools had been receiving all along. Because of the nature of a vocational school, students typically divided their time between vocations and academics. One teacher commented:

Since the exit exam deals [primarily] with English and math specifically, [the principal] doubled up the time when the youngsters in ninth and tenth grade would deal with English and math. So that in itself should give youngsters a little better preparation. We could kind of compensate for the lack of time we usually have. While we still don’t have the same amount of time that others who aren’t vocational schools, have, it’s certainly a huge improvement.

In summary, the structural or technical changes in the school are dramatic as the ninth and tenth grade curriculum aligned more closely to the state-mandated test. When viewing the changes that have occurred in our study school, on the surface, it appeared that these changes had a good chance of being institutionalized and ultimately leading to the goal of raising test scores. It is not yet clear as to what impact these efforts will have on student test scores.

The district superintendent expressed his support of the scheduling changes that came out of the teachers’ and principal’s decision-making:

The five week in vocational and five weeks on academic was something that the teachers have been talking about and couldn’t really decide whether they wanted to do or not. And the sixth period was something that Mr.–[principal] had brought with him from the other school that allowed for more teaming, which was an asset.

The teachers were then asked to vote on these changes and submit recommendations:

The principal and the City and the School Committee has approved that we shift to exit exam assistant courses. So additional read/write in English and additional math are being offered in school. Which we felt, as a school, ‘cause it was discussed with us, for our recommendations. And the principal did present this to us for a vote, asking the faculty if we’d like to go from five to six periods and what would you like to do with the extra period ... We agreed that that would be in the school’s best interest and the students’ best interest ... [The vote itself] was quite overwhelming. I thought it would be a close vote because I felt that some people just didn’t want to do the extra preparation to teach the extra courses. From what I understood with approximately 160 staff members or so, it was like 90 something in the majority. It was a definite positive response that the principal received. It was a mandate from the faculty that “We’re behind you. We want to try to put VHS back on track. If this is what you’re asking us to do, we’ll do it during school hours.

## STATE SUPPORT AND RESOURCES

Overall, the teachers for the most part seemed grateful and relieved that the state had stepped in and taken some sort of action. A teacher said:

In a lot of ways, personally for me, it really has helped me focus more on what kinds of things I can do to help the students. What kind of strategies. How to take data from EE and look at what the students need and focus my teaching generally around some of these things.

The state and the school seemed to have a benevolent yet symbiotic relationship that remained fluid, as together, the state and the school negotiated unfamiliar territory. One teacher said, “Yeah the state, to be honest with you is getting their feet wet in it too. They’re kind of one step ahead of us as we go down the road and they’re taking note of the problems that we have and they’re adjusting accordingly.” What may be important here is that efforts on behalf of the state were not only unidirectional but reciprocal as state and school coordinated their efforts. The state in particular, provided the pressure and resources to help alleviate one school’s problem of low student performance.

Some teachers reported positive effects as a result of the underperforming designation:

I think it’s more on the positive [side] and now we’re trying to get more out of the students ... now we’re more focused on what they really need, their weaknesses and what their strengths are and from there we’ve started basing our curriculum. I think it’s more structured now. Now, knowing that they need more reading skills, that they need more writing skills, that they need more math skills. It’s more structured.

This same teacher continued by saying that as a result, she was more committed to helping the students since she knew better how to help:

I think now I’m more into helping the students. We’re more helping the school. I see myself involved in every way now, and for the good ... I think most of the [teachers] are being very positive ... Everybody’s helping each other now. I think we’re more into ‘Let’s do it.’ More in tune actually than before.

The state introduced a review framework that involved using data to look at specific populations of students to determine what could be done to help them. The district superintendent commended this intervention:

I think it’s a very good direction to do good diagnosis and to look at what the learning gaps are in any school. I think you need to be able to design an effective school improvement plan that really targets what the needs of the students are. So I think that’s very helpful. I mean you have to look at the data to know what the students’ needs are.

This strategy also influenced teachers' own practice of data use. One teacher explained:

What I've done with a colleague of mine that's not in this school but another English teacher, we've gone through and we set up the standards of what was on the SEE and the questions, what questions addressed these particular standards. From there, we looked at our students' individual reports and found that say there were 8 questions per standard A, which was basic facts or main ideas for each student and it's all done manually. Find out how many questions they got wrong. If they got seven out of eight questions for basic facts and main ideas, and they're weak in that area, so you need to go through and build up just some basic reading strategies and comprehension skills. So that's what I've done. It's time consuming to go through and do it, but once you have that information, it is extremely valuable.

Another teacher explained that the school was involved in conducting item analyses of student scores: "Then we do item and data analysis to determine where are we good and where are we bad and where our curriculum could go and where our teaching styles have to hit." The principal summed up the situation at the school by saying:

Well it's sort of been coming on because SEE, for the last four years has really provided data and this issue about reviewing data and making decisions based on data. While it's not new, change comes slowly sometimes and our public school model has historically been to determine change by the seat of our pants ... We haven't gone out and measured kids in a big way ... I really do believe [using data] has a lot of merit.

## SELECTIVITY

The state also allowed VHS to carefully select incoming students.<sup>8</sup> The policy seemed to provide teachers some hope and reprieve with the possibility of raising their test scores. One teacher commented:

I think it's a new day for our building. They talk about the new Vocational HS and the new reality [due to] this incoming class. I believe from my exposure to them, it is going to provide us with what is needed.

As one teacher explained, the school had always been viewed as a 'dumping ground' in the district and now things were changing:

I think that we had been dumped on for years. I mean this school, and I'm not telling any tales on the school 'cause Dr.-[superintendent] would be right here championing this same comment ... Anything that didn't work well in the system cuts in here. Staff members, it could be custodians, it could be whomever. Students. You name it ... They just said, you know, they should be at Vocational HS if they didn't make it anywhere else or if there were problems. So this building was really filled with handicapped kids and



teachers and custodians and everything else ... Things are starting to change. It's an absolutely amazing change.

A previous principal, who was a superintendent in the district at the time of our visit, told us, that the school has come a long way. She described the school when she first became principal:

When I first went to that school, there are no textbooks. The kids never were taking notes. Nobody was teaching the kids. The kids not only had their heads on their desk during the classes, they had their hoods over their heads. Now, there's not a class that you could go into that the kids don't have their notebooks, that they're not paying attention. And most of the teachers are using appropriate instructional models in the classrooms. [I say to myself,] "Would you look at how far those teachers took these kids that everybody else had given up on. It's a miracle. It's a miracle that they don't even walk out of here because they're so exhausted and then somebody says you did a bad job 'cause they didn't pass the exit exam.

## STUDENTS' VIEWS

It is important to note that these students we spoke with passed the EE and did not face the same challenges that their failing peers did. However, they were able to speak from their perspective of knowing these other individuals, which we have included as anecdotal data. One vocational teacher expressed the challenges and disappointments many students faced that graduating year:

[This crisis is] a reality check for [students] certainly. Because somebody outside of [this city] has told them that fifty percent of their education doesn't matter. The time they spend in their vocational time, half of their time, is not what matters. That whatever they do up here, in their academics drives their career. And like I said, when these kids drop out of school, they don't drop out because they didn't do well in the vocational. They drop out because they struggle to do their academics. So it was a reality check that life isn't always fair and that they just have to move on the best they can and that's it. We have, I don't know if you heard the numbers or not, but my understanding is that about 15% of the graduating class of 2003 [will receive their diploma]. The other 85% aren't because they didn't pass the exit exam.

The reality of the EE became starkly evident to us from the words of students who had peers who failed. One student commented:

So you can't leave until you pass EE, and I think that's really stupid because a lot of kids are still in school because of the EE; they're all seniors and they can't get out 'til they pass the EE. And they're just gonna drop out because of that. Now they're all ready to get out but then if they don't graduate, they can't leave.

The severity of the situation described above becomes crystallized when we realized how committed one of the students was about her career since she chose VHS specifically because of the vocational training it provided:

Well, when I get to twelfth grade, I'm gonna take my state boards and then I'll get my license [in cosmetology] if I pass. You have to get to your license. And then pretty much I can do whatever I want from there. But I want to have my own business so I gotta go to college and take business management courses for that.

This same student revealed to us that she had some understanding of the testing process that included having the opportunity to retake the EE, yet she also noted that by not passing the EE, students did not receive their graduation diploma.

I think you can get to take it [EE] over again until you pass it and if you don't pass it and you're in twelfth grade, you just don't get your diploma ... and your certificate that says you went to school here ... I don't think you can really do anything with that ... unless you pass.

Two of the four students we spoke with questioned why one test should determine their high school graduation:

We can pass the ninth into tenth and eleventh grade; we don't have to take a test to get out. So why should we have to take a test to graduate?

After 12 years of doing everything, passing all your classes, it's down to one test?

## **PERSISTING INEQUALITIES**

As we listened to what the teachers and students were saying about this experience, we realized that some critical and valid concerns were emerging regarding the testing policy the teachers in our study school had been asked to implement. The most salient arguments revolved around notions of equity and fairness. Several attributes of VHS repeatedly arose which contributed to low student achievement. Here's a sketch of these reported constraints:

High special education and bilingual population:

I think we're at a disadvantage because we're an urban school because of our high special ed and bilingual population. I think the students are coming in here because of being dumped into a vocational school. The foundation skills are not there. As far as how it's affecting the students, it's very stressful for them.

Bilingual students, like every other student, have to pass the EE, and it has to be passed in English if they wish to have a high school diploma.

### Worries over student engagement:

I think you have to deal with the value that they have for education. If they don't value it then it's not going to mean anything to them and unless it means something, they're not going to put the effort in that's going to be necessary. Gangs in the cities have more influence by far than the school does ... And those are not just mere distractions, they're one's life. Social security. How they get money from the household. All of these financial issues, cultural, attitudinal, value, are just overwhelming when it comes to trying to get a youngster to do something as divorced from his values as taking EE is.

### VHS students arrive under prepared, so middle schools must be held accountable:

The [state] is talking about what can we do with a high schooler ... that comes to the school without knowing how to do fractions and how to read. [My question is this:] Why is the school under performing when we picked up the students to come over here to the school from some other place, from some other schools? So if there are some schools that should be under performing, they should be the schools where those particular students were before they come over here.

We have students here who ... come here and they are still reading on the second, third, fourth, fifth grade level. Well what happened to those children and you know, how did the system fail them? And how did all of a sudden we now become responsible for remediating the last eight or nine years of their education?

### Poor student attitudes:

The majority of the kids, I think, have a negative feeling about it. They're scared. They think they're not prepared. They don't think they can do as well as other schools, non-vocational schools, non-inner city schools.

They look at it like, "I don't want to take it. I'm not going to take it. I'm not going to do well on it. They just [say], "I can't pass that ... They just automatically say, "I can't do it." I mean a lot of these kids I know they feel they can't.

Of course there are some who feel, "I'll never pass the EE." And those are the students that will drop out unfortunately.

## **DISCUSSION – THE DILEMMAS OF HIGH STAKES TESTING**

Systemic reform efforts are generally based on the assumption that public schools represent rational organizations that can make structural changes to accommodate policy shifts. We did observe at VHS a variety of structural changes, including a new schedule, timetable, and curricular programming. This type of examination depicts actions and decisions that

are unidirectional and technical in nature as the focus on a single desired goal of raising test scores.

But, when we began asking deeper questions such as: Who determines these changes? What is the impact of these changes? We begin to see more enduring and in some ways, troubling facets of standardized testing and its impact on schooling. While on the surface, raising test scores may appear to be relatively straightforward, such an endeavor actually has significant implications for the individuals who make up the school. “The agency of educators is part of a complex dynamic, [that] shap[es] and [is] shaped by the structural and cultural features of school and society” (Datnow, et al., 2002, pp. 12–13). In this way, we must recognize the context in which policy implementation takes place, the “face-to-face interactions among real people, confronting real problems in concrete social contexts, such as classrooms, school board meetings, courts and other contexts.”

While our case study cannot be generalized, we are able to draw some tentative conclusions. From a traditional policy analysis perspective, the structural changes in the school were dramatic as the ninth and tenth grade curriculum became aligned more closely to the state-mandated test. The changes were comprised of “systematic procedures, the sequencing of activities, or other procedural inventions [that were] designed to solve problems that stand in the way of organizational task achievement” (Owens, 2001, p. 93). Due to the number of scheduling and curricular changes that occurred, there were positive and negative consequences for teachers and students. The most noticeable change was the structure and programming and the emphasis on academics. Indeed, VHS’s exit exam performance has improved each year since it became a graduation requirement, although the school continues to lag behind most other schools in the district and state. The obvious negative was the reduction of the vocational mandate of this vocational school, as teachers’ daily lives revolved around the single goal of raising student academic achievement levels.

Secondly, we dug deep to examine how these changes played out in the daily lives of educators and students inside one school that experienced a swirl of reform pressures. To our surprise, there was little to no resistance to these changes on the part of the educators we interviewed. In cases such as this, often communities of teachers are mobilized to respond to policy pressure (Datnow et al., 2002), motivated to problem-solve the strategies needed to meet their local needs (Berman & McLaughlin, 1978), and come to a consensus about whether the changes are both desirable and doable in

relation to the desired goals (Fullan, 1997). While these characteristics are often depicted as signs of successful policy implementation, it would be unrealistic to claim that teachers were not part of a 'culture of compliance' (Blackmore, 1998) as they strove to implement a policy that was both mandated and sanctioned. Despite these expectations for conformity and complicity, however, the teachers for the most part seemed grateful and relieved that the state had stepped in and taken some sort of action.

Academic teachers were engaged in careful analysis of test scores to gain an understanding of their students' weaknesses. However, the vocational teachers struggled with the increased emphasis of academics in their shops. The majority of teachers expressed concerns about the loss of a vocational focus. The changes introduced and implemented at VHS seemed to convey that state intervention had brought with it a formal structure in which the faculty could address their past problems. Consequently, the state and the school worked cooperatively to address issues of concern. The school reaped the benefits of a review framework the state introduced that involved using data to analyze achievement patterns. It became evident, at least on the surface that VHS was benefiting from state intervention. While the school had taken a pronounced technical interest in addressing the new accountability policy in its efforts to raise student achievement, they now had a united goal and new tools with which to improve their practices.

The efforts on behalf of the state were not only unidirectional but reciprocal as state and school coordinated their efforts. The state in particular, provided the pressure and resources to help alleviate one school's problem of low student performance. An important factor needed to sustain a policy is the institutional capacity to carry out reforms by state officials and those toward whom the policies are directed such as school districts, local schools, teachers, parents, or students. Yet it is, in reality, a significant factor standing between policy adoptions. It is not merely whether state department of education bureaucrats wish for a policy to succeed or fail; it is whether there is organizational capacity sufficient to support successful implementation (Cibulka & Derlin, 1998, p. 504).

VHS's experience with its state exit exam leads to serious questions regarding the relevance of a high stakes test in a school that's earnestly trying to engage students via academic and vocational approaches. While the teachers in our study were not necessarily opposed to testing or accountability, they did express concerns about the nature of the tests and how they were being used (Popham, 2000; Earl, 1999). Studies such as this one urge policymakers to take a broad view of the problem of high stakes

tests and of possible solutions. “Testing can be an important ingredient in a reform initiative, but it constitutes only one of a number of necessary policy changes. Testing certainly is not and cannot be the main engine of school reform.” (McLaughlin, 1991, p. 251)

This kind of in-depth case study illuminates particular and diverse contexts of schooling into which high stakes testing is being installed. We present a story of one school suggesting that policymakers must continue their examination of the dilemmas, organizational responses and mixed effects which stem from the policy shock represented by exit exams especially when high-stakes testing of academic skills conflicts so sharply with a school’s institutional history.

## NOTES

1. Data have been rounded in order to ensure anonymity.

2. This is the third stage in the process used to assess school performance in a state’s accountability system and is not included in the references in order to maintain anonymity for the state, district and school.

3. An alternative district or state credential honoring other achievements was given to about 200 students who appealed the EE. In addition, 300 students were granted waivers or certificates of attainment if they had taken the EE three times and scored within four point of the passing rate, that they had a near perfect attendance, that their grades compared to students who passed the EE, and they had a teacher recommendation. Additionally, 449 students who had failed the EE in 2003 learned that they had actually passed and were able to graduate that year when State Education Officials were convinced by a student to accept an alternative answer to an EE math question. The alternatives mentioned above expired in 2005 (School Panel Review Report, 2002).

4. For the purpose of our larger study, disadvantaged students are defined as those students living in poverty (indexed by participation in the federal free and reduced-price lunch program), and those who are members of groups that have been historically discriminated against in U.S. society (Castellano et al., 2002). CTE refers to career and technical education, previously called vocational education.

5. One of which serves as the study school discussed in this chapter.

6. The vocational programs in our study school were called shops, and we use this terminology.

7. The use of quotes by the participants are intended to provide more of a narrative of what the inhabitants of the school and district were feeling, valuing, and experiencing during this time of crisis. The data are rich in their story-telling value but lack rigor in their (in)ability to provide a complex picture of varying beliefs or values, primarily due to the limited number of responses to the topic.

8. It remained unclear to us if this is a direct result of the school being declared underperforming or if this policy is simply a fortuitous policy that coincided with the all that was going on at our study school.

## ACKNOWLEDGMENTS

The work reported herein was supported under the National Research Center for Career and Technical Education Program, PR/Award No.VO51A990006 administered by the Office of Vocational and Adult Education, U.S. Department of Education. However, the contents do not necessarily represent the positions or policies of the Office of Vocational and Adult Education or the U.S. Department of Education, and you should not assume endorsement by the Federal Government.

## REFERENCES

- Austin, J., & Mahlman, R. A. (2002). *High-stakes testing: Implications for career and technical education* (<http://icdl.uncg.edu/ft/0604030-01.html>). Columbus: OH, National Dissemination Center for Career and Technical Education.
- Ball, S. J. (1994). *Education reform: A critical post-structural approach*. Buckingham: Open University Press.
- Berman, P., & McLaughlin, M. W. (1978). *Federal programs supporting educational change* (Vol. VIII). Santa Monica, CA: RAND.
- Blackmore, J. (1998). The politics of gender and educational change: Managing gender of changing gender relations? In: A. Hargreaves, A. Lieberman, M. Fullan & D. Hopkins (Eds), *International handbook of educational change* (pp. 460–481). Norwell, MA: Kluwer Academic.
- Capper, C. A. (Ed.) (1993). *Educational administration in a pluralistic society*. NY: State University of New York Press.
- Castellano, M., Stringfield, S., & Stone, J. R., III. (2002). *Helping disadvantaged youth succeed in school: Second-year findings from a longitudinal study of CTE-based whole-school reforms* (Retrieved on September 9, 2005 from <http://www.nccte.com/publications/secure/index.asp#>). Minneapolis: University of Minnesota, National Research Center for Career and Technical Education.
- Castruita, R.M. (2006). Why exit exam decision is a good one, June 2. *The San Diego union tribune*. [www.signonsandiego.com/uniontrib/20060602/news\\_lz1e2castruit.html](http://www.signonsandiego.com/uniontrib/20060602/news_lz1e2castruit.html). Retrieved on October 8, 2007.
- Center on Education Policy, June. (2005). [www.ecs.org/html/IssueSection.asp?issueid=108&subissueid=159&s=Selected+Research+%26+Readings](http://www.ecs.org/html/IssueSection.asp?issueid=108&subissueid=159&s=Selected+Research+%26+Readings). [http://www.nsba.org/site/doc\\_sbn.asp?TRACKID=&VID=58&CID=810&DID=10901](http://www.nsba.org/site/doc_sbn.asp?TRACKID=&VID=58&CID=810&DID=10901). Retrieved on October 8, 2007.
- Cibulka, J. G., & Derlin, R. L. (1998). Accountability policy adoption to policy sustainability: Reforms and systemic initiatives in Colorado and Maryland. *Education and Urban Society*, 30(4), 502–515.
- CNN.com/Education. (2003). States stick with high-school exit exams, August 20. <http://www.cnn.com/2003/EDUCATION/08/13/high.school.exams.ap/index.html>. Retrieved on October 8, 2007.

- Datnow, A., Hubbard, L., & Mehan, B. (2002). *Extending educational reform: From one school to many*. London: Falmer.
- Earl, L. (1999). Assessment and accountability in education: Improvement or surveillance? *Education Canada*, 39(3), 4–6, 47.
- Earl, L. M., & LeMaheiu, P. G. (1997). Rethinking assessment and accountability. In: A. Hargreaves (Ed.), *ASCD yearbook: Rethinking educational change with heart and mind*. Virginia: ASCD.
- FairTest Examiner. (2007). Exit exam update: MD, WA, MA, CA, PA, April. <http://www.fairtest.org/examarts/2007%20April/ExitExams.html>. Retrieved on October 8, 2007.
- Firestone, W. A. (1998). A tale of two tests: Tensions in assessment policy. *Assessment in Education*, 5(2), 175–191.
- Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972–1977*. Brighton: Harvester Press.
- Fullan, M. (1997). Constructive change for complex times. In: A. Hargreaves (Ed.), *Positive change for school success*. Alexandria, VA: ASCD Yearbook.
- Greene, J. P., & Winters, M. A. (2004). High school exit exams don't cause dropouts. *The San Diego Union-Tribune*, May 19. [http://www.signonsandiego.com/uniontrib/20040519/news\\_lz1e19greene.html](http://www.signonsandiego.com/uniontrib/20040519/news_lz1e19greene.html). Retrieved on October 8, 2007.
- Hamel, J., Dufour, S., & Fortin, D. (1993). *Case study methods*. Newbury Park, CA: Sage Publications.
- Hargreaves, A., Earl, L., & Schmidt, M. (2002). Perspectives on alternative assessment reform. *American Educational Research Journal*, 39(1), 69–95.
- Kemple, J. J., & Snipes, J. C. (2000). *Career academies: Impacts on students' engagement and performance in high school*. New York: Manpower Demonstration Research Corporation.
- Lynch, R. L. (2002). High school career and technical education for the first decade of the 21st century. *Journal of Vocational Education Research*, 25(2), 155–198.
- MacPherson, R. J. S. (1996). Educative accountability policy research, theory and practice. Paper presented to the Office of Educational Quality and Accountability Office, Toronto, Ontario.
- McGinley, S. (2002). *High-Stakes testing isn't the answer: Study compares AIMS test scores for different student populations*. <http://ag.arizona.edu/pubs/general/resrpt2002/13.pdf>
- McLaughlin, M. (1991). Test-based accountability as a reform strategy. *Phi Delta Kappan*, 73(3), 248–251.
- Mendez, T. (2004). *Exit exams yield a graduation gap*. The Christian Science Monitor, August 19. <http://www.csmonitor.com/2004/0819/p12s01-legn.html>. Retrieved on October 8, 2007.
- Owens, R. G. (2001). *Organizational behavior in education: Instructional leadership and school reform*. Boston, MA: Allyn & Bacon.
- Peterson, K. (2005). *High school exit exams on the rise*. Stateline.org, May 23. <http://www.stateline.org/live/ViewPage.action?siteNodeld = = 136&languageId = 1&contentId = 33244>. Retrieved on October 8, 2007.
- Popham, W. J. (2000). The mismeasurement of educational quality. *School Administrator*, 57(11), 12–15.
- Scheurich, J. J. (1994). Policy archeology: A new policy studies methodology. *Journal of Education Policy*, 9(4), 297–316.
- Sirotnick, K., & Oakes, J. (Eds). (1986). *Critical perspectives on the organization and improvement and schooling*. Boston: Kluwer-Nejhoff.



- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park: Sage.
- Taylor, S., Rizvi, F., Lingard, B., & Henry, M. (1997). *Educational Policy and the Politics of Change*. New York: RoutledgeFalmer.
- Yin, R. (1993). *Applications of case study research*. Beverly Hills, CA: Sage Publishing.

## ABOUT THE AUTHORS

**Soung Bae** is a doctoral candidate in the Human Development and Education program at the University of California, Berkeley. Her work examines how teacher learning and development is best cultivated and the relation between beliefs and practice.

**Marisa Castellano** is Visiting Associate Professor in the College of Education and Human Development at the University of Louisville. Her research interests span K-14, and include identifying strategies that can improve the opportunities of minority students.

**Gayle S. Christensen** is the Associate Director of International Initiatives and Projects in the Office of the Provost at the University of Pennsylvania. Her cross-national research has focused on two main areas: immigrant student achievement and accountability.

**Bruce Fuller** is Professor of Education and Public Policy at the University of California, Berkeley. His work focuses on how policy makers attempt to touch the daily activities of educators and children, as examined in his new book, *Standardized Childhood* (Stanford).

**Kristin Gordon** is a doctoral student in Sociology at Emory University. Her work examines how recent education reform affects the work lives of public school educators and administrators.

**Laura S. Hamilton** is a Senior Behavioral Scientist at the RAND Corporation. Her current research addresses standards-based accountability policies, educator pay for performance, school reform implementation, and school leadership.

**Melissa K. Henne** is a doctoral student in the Policy, Organization, Measurement, and Evaluation program at the U.C. Berkeley Graduate School of Education. Her research interests are school reform, assessment, and accountability.

**Jennifer Jellison Holme** is an Assistant Professor of Educational Policy and Planning at the University of Texas at Austin. Her work centers on the politics and implementation of educational policy, focusing on the relationship between school reform, equity, and diversity.

**Heeju Jang** is a doctoral student in the Policy, Organization, Measurement, and Evaluation program at the U.C. Berkeley Graduate School of Education. Her research interests are teacher education and certification, assessment, and quantitative methods.

**Thomas F. Luschei** is an Assistant Professor of Education Policy at the Learning Systems Institute, Florida State University. His research interests include accountability reform abroad and in the United States, comparative education policy, and teacher labor markets.

**Julie A. Marsh** is a Policy Researcher at the RAND Corporation, who specializes in research on policy implementation, district-level educational reform, accountability, and school-community collaboration.

**Jeremy Miles** is a Quantitative Psychologist at the RAND Corporation; his primary research involves applying and evaluating multivariate modeling techniques.

**Jennifer Lin Russell** is an assistant professor of education policy at the University of Pittsburgh and a research scientist at the Learning Research and Development Center (LRDC). Her research uses the conceptual tools of organizational sociology to explore policy implementation, education organizations/systems, and teachers' work.

**Michèle Schmidt** is Assistant Professor and Associate Director of Graduate Programs in the Faculty of Education at Simon Fraser University. Her work focuses on policy implications of classroom assessment and accountability, social capital in schools, and the emotions of teaching and learning.

**Brian M. Stecher** is a Senior Social Scientist at the RAND Corporation. His work focuses on assessment and accountability in education with a particular emphasis on their impact on classroom practice.

**Dr James R. Stone, III** is a Distinguished University Professor at the University of Louisville. He has been the Director of the National Research Center for Career & Technical Education since 2002.

**Sam Stringfield** is a Distinguished University Professor at the University of Louisville. His research focuses on strategies for improving the academic achievements of disadvantaged populations in high-poverty schools.

**Athena Tapales** is currently an independent scholar. Most recently she was a post-doctoral fellow at Johns Hopkins University.

# SUBJECT INDEX

- Academic Performance Index (API), 81, 138–139
- Accountability
  - and coupling of schools, 10–12
  - federalist model, 14
  - and NCLB Act, 35–36
  - organizational mechanisms and local effects of top-down, 8
  - school's internal, 71
  - standards-based, 11
  - and working condition, 107–109
- Adequate Yearly Progress (AYP) targets, analysis, 22, 43, 49–50, 55–58, 60, 104, 113–114
- Advancement Via Individual Determination (AVID) program philosophy, 172
  
- Bae, Soung, 23, 189
- Borko, H., 19
- Bush, George H. W., 5, 7
  
- California, accountability system in data methods
  - classification of schools and districts, 79–80
  - rating of schools and districts, 77–79
  - sampling strategy and data, 76–77
  - weak and strong accountability, 76
- discussion, 90–98
- results, 81
  - state accountability context, 80–81
  - Urban district, 88–89
  - Vine district, 89–90
- theoretical framework
  - internal accountability and loose coupling, 71–72
  - responses to accountability policies, 72–75
  - role of district, 69–71
- California Achievement Test, 6th Edition (CAT6) tests, 81
- California High School Exit Exam (CAHSEE), 80–84, 86, 88–89, 162–183, 187–188
- California's Public School Accountability Act (PSAA), 139
- Capacity building, 161–162
- Castellano, Marisa, 209
- Centralized curricula and high-stakes testing, 6–7
- Centralized environments and accountability policies, 137
- Christensen, Gayle, 22, 67
- Chubb, John, 10
- Clinton, Bill, 4–5, 9
- Cohen, Michael, 10–12
- Coupling, of education system
  - internal accountability, 71–72
  - and policy responses, 72–75
- Cross, Christopher, 5
  
- Decentralized environment and accountability policies, 135–137
- District's role, in accountability reform data methods
  - classification of schools and districts, 79–80

- rating of schools and districts, 77–79
  - sampling strategy and data, 76–77
  - weak and strong accountability, 76
- discussion, 90–98
- results, 81
  - relationship between district efforts and schools' responses, 94–98
  - state accountability context, 80–81
  - variation in responses, 93–94
- theoretical framework
  - internal accountability and loose coupling, 71–72
  - responses to accountability policies, 72–75
  - role of district, 69–71
- District's role, in teacher learning and development. *See also* California, accountability system in
  - coherence, 70–71
  - discussion, 203–205
  - findings
    - coaching model, 197–199
    - district-wide professional development, 195–197
    - human capital, 194–195
    - leadership, 201–203
    - trust, concept of, 199–201
  - internal accountability, 79–80
  - methods
    - data analysis, 193
    - data collection, 192–193
    - district context, 192
  - theoretical framework, 190–191
- Effective district- or school-level practices, analysis
  - in centralized environments, 137
  - in decentralized environments, 135–136
- discussion and policy implications, 151–154
- findings, 143–151
- indices for, 140–143
- limitations of research, 138
- school sample, 138–140
- Elliott, R., 19
- Empowerment, 108–109, 111, 113–114, 120–123
- Exit examination, in high school
  - background and history, 159–160
  - policy assumptions, 161–162
  - research on high schools' response
    - curriculum change, effect of, 167–173
    - findings of study, 164–167
    - instructional change, impact, 173–183
    - sample schools, 163–164
- Exit examination, in VHS
  - discussion, 228–231
  - estimates, 211–212
  - findings
    - academic achievement, 218
    - challenges and disappointments of students, 226–227
    - inequalities, 227–228
    - reform efforts, 222–223
    - schedule changes and its impact, 219–220
    - school district and VHS, 216–218
    - state support and resources, 224–225
    - students' work experience, 222
    - teacher experiences, 220–222
  - methodology
    - data collection and analysis, 215–216
    - sample, 215
- Exit examination policies, 161–162
- External accountability policies and teachers, 108

- Georgia Quality Learning and Teaching Environments Initiative (QLTE) survey, 109–110
- Gordon, Kristen, 22–23, 103
- Hamilton, L., 22, 31
- Henne, Melissa, 23, 26, 133
- High Priority Schools Grant Program (HPSGP), for schools, 81
- High school response, to exit exam. *See* Exit examination, in high school
- Holme, Jennifer Jellison, 23, 157
- Immediate Intervention/  
Underperforming Schools Program (II/USP), 81
- Implementing Standards-Based Accountability (ISBA) project  
data and methods  
analytical approach, 41  
surveys, 40–41  
discussion, 60–63  
perspectives on organizational reform, 33–34  
practical application, 36–38  
results  
development of a school and district improvement plan, 50–53  
instructional practices related to SBA, 42–50  
relationship between school- and classroom-level responses to SBA, 53–60  
theoretical aspects, 34–36
- Incentives, for accountability, 35
- Instructional leadership, 55
- Instructional practices  
alignment of curriculum and student assessments, 43–45  
attributed to testing, 45–47  
elementary teachers and, 42–43, 46–49  
in Georgia, 49–50  
responses of middle school teachers, 46, 48–49  
role of occupational norms and workplace conditions, 33  
school-level efforts, 34  
time, 42–43  
use of achievement data for instructional decision-making, 47–50
- Internal accountability, 71–72
- Jang, Heeju, 23, 26, 133
- Luschei, Thomas, 22, 67
- Marsh, Julie A., 31
- McNeil, Linda, 18
- Meyer, John W., 10
- Miles, Jeremy, 31
- Minimum competency tests (MCTs), 159
- Mintrop, H., 19
- Moe, Terry, 10
- No Child Left Behind Act (NCLB), 2001, 7, 31, 50, 67, 104, 108–109, 161
- O’Connell, Jack, 160
- Organizational reform, perspectives on, 33–34
- Porter, Roger, 5
- Reagan, Ronald, 1
- Romberg, T., 19
- Russell, Jennifer Lin, 31
- Schmidt, Michèle, 23–24, 209
- School- and classroom-level responses to SBA, analysis  
instructional leadership activity measures, 55

- multilevel modeling approach, 56–60
  - principal measures, 54–55
  - results for teacher's use of
    - achievement data, 59–60
  - teacher measures, 55–56
  - test-focused instructional practices, 57–58
- Similar Students, Differing Results (SSDR) study, 135
- Smith, Marshall, 10
- Specially Designed Academic Instruction in English (SDAIE), 87
- Standards-based accountability (SBA), 11
  - action mechanisms, 36
  - direct changes in teacher practice, 37
  - effects at the school level, 37–38
  - instructional practices
    - alignment with curriculum and student assessments, 43–45
    - attributed to testing, 45–47
    - time, 42–43
    - use of achievement data for instructional decision-making, 47–50
  - in practice, 36–38
  - student learning, impacts, 37
  - in theory, 34–36
  - under NCLB Act, 35–36, 50
- Stanford Achievement Test, 9th Edition (SAT9), 81, 88–89
- Stecher, Brian, 19, 31
- Stone, James R., III, 209
- Stringfield, Sam, 209
- Student achievement
  - and accountability policies, 6–14
  - centralization of education, 4–6
  - and government policy, 2–4
  - relationship with teacher-reported practices
    - district staff emphasize on student learning, 141
    - and ethnic groups, 143–151
    - focus on school's curriculum, 142
    - focus on student subgroups, 141–142
    - modification of pedagogical practices, 141
    - policies regarding student behavior, 141
    - positive environment, 143
    - school focus on, 140–142
    - school priorities, 141–142
    - teacher development for serving English learners, 142
    - teacher expectations, 141
    - teachers use of curricular standards, 140
    - in United States, 1
- Tapales, Athena, 209
- Teaching practices and accountability
  - data and methods
    - analytical approach, 41
    - surveys, 40–41
  - discussion, 60–63
  - perspectives on organizational reform, 33–34
  - practical application, 36–38
  - results
    - consistency of teachers' responses, 50–53
    - development of a school and district improvement plan, 50–53
    - instructional practices related to SBA, 42–50
    - relationship between school- and classroom-level responses to SBA, 53–60
    - theoretical aspects, 34–36
- Thatcher, Margaret, 1, 11–12



- Top-down accountability policies,  
105–107, 123–126
- Trujillo, T., 19
- U.S. job growth and educational  
system, 160
- Valenzuela, Angela, 18
- Vocational high school response, to exit  
exam. *See* Exit examination, in VHS
- Williams, S., 19
- Woody, Elisabeth, 19
- Working condition  
and accountability,  
107–109  
teachers' assessments of,  
109–114  
qualitative analysis, 114–123
- Zarinia, E., 19