RESEARCH FOR *ACTION*

An Occasional Paper

Complicated Choices:
Struggling to Meet NCLB Requirements AND Remain Faithful to a School's Educational Vision and Practice

NOVEMBER 2009

Complicated Choices:
Struggling to Meet NCLB Requirements AND Remain Faithful to a School's Educational Vision and Practice

Prepared by Research for Action

Eva Travers

An Occasional Paper

Table of Contents

Introduction

In this era of high-stakes accountability, school leaders often face contradictory pressures as they strive to improve student performance. They must meet the federal mandate of No Child Left Behind (NCLB) for student achievement in tested subjects on a yearly basis. At the same time, many believe that NCLB constrains their professional judgment about how to best teach and assess students in the context of their own schools.

The following case study illustrates how the current accountability system can lead to inequity because of the disincentives to be responsive to the educational needs of all students and the complicated choices this can force educators to make. As the reauthorization of NCLB proceeds, it is important to consider the diversity of American schools, and students attending them, in order to design an accountability system that is more equitable and responsive. In this process, the experiences and practices of particular schools can offer valuable lessons.

Between 2005 and 2007 I conducted field research at Baker,[1] a K-8 school in Philadelphia. From my perspective, this was a school that used a variety of exemplary leadership and faculty development practices and whose educational goals and approaches supported adult and student learning. Longitudinal data showed the school had made significant progress in student achievement, yet the school had failed to make Adequate Yearly Progress (AYP) according to the NCLB accountability metric.

The school made complicated choices as it attempted to balance NCLB requirements AND remain true to its educational vision and local practices.

This paper examines the complicated choices the school made as it attempted to balance NCLB requirements *and* remain true to its educational vision and local practices. It also discusses the pedagogical and psychological consequences of NCLB for the school. In the Conclusion, I suggest some policy implications of the Baker case.

My initial impressions of Baker, located in one of the most impoverished neighborhoods in Philadelphia remain vivid. Examples of student art—from papier-mâché animals and elaborate woodcuts to watercolors of landscapes and painted tiles—filled the open landing next to the Main Office, the "Baker museum," as it is called. Faculty and staff invariably greeted me with friendly smiles and made me feel welcome in the school. The halls were quiet and student work posted on the walls outside many classrooms suggested that effective teaching and learning were occurring inside them.

The student body at Baker, approximately 650-700 pupils during the time of my research, was relatively diverse. In the 2006-2007 school year, 84 percent of the student body was African American, 8 percent was Latino, and 7 percent was Asian. Fourteen percent received Special Education services and 9 percent received ESOL instruction, a relatively

---

[1] A pseudonym.

high number of special needs students for a K-8 school in the district. Eighty-five percent received free or reduced lunch.[2]

To be candid, my early positive impressions of the school, educational philosophy, practice as a progressive teacher educator, and firsthand appreciation of the challenges of teaching and learning in urban schools framed the way I looked at Baker throughout my research.

Early in my fieldwork, I observed that Baker was implementing "best practices" in a number of areas, such as:

- Distributed leadership,[3]
- Teacher collaboration,[4]
- Instructional coaching, and
- Professional development geared to the specific goals of the school.[5]

I also learned that Baker's goals included student-centered and differentiated instruction[6] and that the school emphasized performance-based assessment, writing, and higher-order thinking in order to improve individual student learning.[7] While Baker was characterized as a failing school by NCLB's accountability metric, this description contradicted my

---

[2] The district average of students receiving subsidized lunch was 74 percent. All statistics from: Demographics/Data, School Profiles. School District of Philadelphia, 2006.

[3] Elmore, R. (2000). *Building a new structure for school leadership.* Washington, DC: The Albert Shanker Institute; Lambert, L. (2002). A framework for shared leadership. *Educational Leadership, 59*(8), 37-40; Spillane J., Halverson, D. & Diamond, J. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23-28.

[4] Darling-Hammond, L. & McLaughlin, M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(8), 597–604; Fullan, M. & Hargreaves, A. (1991). *What's worth fighting for in your school?* New York: Teachers College Press; Inger, M. (1993). Teacher collaboration in secondary schools. *Center Focus, 2;* Lieberman, A. (1990). *Schools as collaborative cultures: Creating the future now.* Bristol, PA: The Falmer Press; Little, J.W. (1987). Teachers as colleagues. In V. Richardson-Koehler (Ed.), *Educators' handbook: A research perspective.* New York: Longman.

[5] Annenberg Institute for School Reform. (2004*). Instructional coaching: Professional development strategies that work.* Providence, RI: Brown University; Herlihy, C. & Kemple, J. (2004). *The Talent Development middle school model: context, components, and initial impacts on students' performance and attendance.* New York: MDRC; Neufeld, B. & Roper, D. (2003*). Coaching: A strategy for developing instructional capacity, promises, and practicalities.* Providence, RI: Annenberg Institute for School Reform; Poglinco, S., Bach, A., Hovde, K., Rosenblum, S., Saunders, M. & Supovitz, J. (2003). *The heart of the matter: The coaching model in America's Choice schools.* Philadelphia, PA: Consortium for Policy Research in Education; Western Regional Educational Laboratory. (2000). *Teachers who learn, kids who achieve: A look at schools with model professional development.* San Francisco, CA: West Ed.

[6] Bravmann, S. (2004).Two, four, six, eight, let's all differentiate! Differential education: Yesterday, today and tomorrow. *New Horizons for Learning*; Laurence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefit the whole class. *American Secondary Education, 32*(3), 34-62; Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners.* Association for Supervision and Curriculum Development. Tomlinson, C. (2004). *How to differentiate instruction in mixed-ability classrooms.* Upper Saddle River, NJ: Prentice Hall.

[7] Coalition of Essential Schools www.essentialschools.com; Darling-Hammond, L., Ancess, J. & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press; Stiggins, R. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement, Issues and Practice, 20*(3), 5-15; Wiggins, C. (1998). *Educational assessment: Designing assessments to inform and interpret student performance.* San Francisco, CA: Jossey-Bass.

initial impressions of the school and subsequent research. Could my professional judgment be so mistaken?

This paper is an attempt to make sense of what I saw happening at Baker, a school that impressed me more and more over the two years that I spent there as an increasingly "engaged researcher."[8]

What follows is the story of a school that

- Struggled to maintain its progressive educational philosophy and practice in the face of NCLB's requirement to achieve Adequate Yearly Progress (AYP);
- Adhered to the School District of Philadelphia's highly prescriptive approach to curriculum and instruction using a Managed Instruction System, while simultaneously continuing to implement local practices that were congruent with a pedagogical identity developed during previous reforms at the school;
- Built a successful learning community for adults and students whose significant accomplishments were obscured by the fact that it did not measure up according to NCLB proficiency requirements. If the U.S. Department of Education had permitted the state to use its valued-added growth model, the Pennsylvania Value-Added Assessment, as an alternative measure of accountability, however, it would have made AYP.

Two Perspectives on Baker's Academic Performance

Initially, the only metric NCLB allowed to measure progress toward its goal that 100 percent of students would achieve proficiency by 2014 was a "status achievement model." [9] This model requires schools to increase the percentage of proficient students on a state test in a step-wise fashion. If schools do not meet the absolute proficiency targets, AYP can be met by demonstrating improvement. However, this improvement is determined by comparing student performance, both overall and by different subgroups, within the same grade and school from year to year. For instance, achievement of

---

[8] Describing myself as an "engaged researcher" fits within the methodology of engaged anthropology/ ethnography that guides activist scholars whose experiential, contextualized research contributes to social change. It is an approach espoused by many feminist scholars that draws upon female tendencies toward empathy and concern and seeks to establish reciprocal relationships with individuals whose stories they are telling. See Lather, P. (1991*). Getting smart: Feminist research and pedagogy with/in the postmodern.* London: Routledge, Chapman and Hall, Inc.; Stacey, J. (1998). *Can there be a feminist ethnography?* Amsterdam: Elsevier Science, Ltd.

[9] Every state is allowed to determine its own proficiency standards as well as the tests it uses, which undermines comparability between states on AYP outcomes. See Jacobs, E. (January 13, 2009). NCLB: Recipe for success or failure? Washington, DC: *D.C. Examiner*; Lee, J. (2008). Two takes on the impact of NCLB on academic achievement. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; Linn, R. (2003). Accountability, responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3-13; Linn, R. (2008). Toward a more effective definition of AYP. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; Marion, S., White, C., Carlson, D., Erpenbach, W., Rabinowitz, A. & Sheinker, J. (2002). Making valid and reliable decisions in determining AYP. *Implementing the state accountability requirements under the No Child Left Behind Act of 2001.* Washington, DC: Council of Chief State School Officers; Mintrop, H. (2008). Low-performing schools' programs and state capacity requirements. In G. Sunderman, (Ed.), *Holding NCLB Accountable: Achieving Equity, Accountability and School Reform.* Thousand Oaks, CA: Corwin Press.

subgroups in third grade in one year is compared to these subgroups' achievement in third grade the subsequent year, as opposed to the achievement of these same students in fourth grade.

In Pennsylvania, schools' AYP status is determined by the percentage of students scoring proficient or advanced in reading and math on the Pennsylvania System of Student Assessment (PSSA). In 2006-07, according to NCLB, Baker was in Corrective Action II, 4th year. Schools in Corrective Action face increasing sanctions every year they do not achieve AYP. The student performance on the PSSA at Baker had never been high enough to meet the state's absolute proficiency targets to make AYP.[10]

Because Baker had not achieved AYP in any year, the school faced possible sanctions.

By 2006-07, Baker's scores had improved substantially, but they continued to be below the state's proficiency targets. For example, between 2002 and 2007 the percentage of eighth graders scoring "advanced" or "proficient" in math improved from 12 percent to 34 percent but the percentage of students required to achieve AYP in 2007 was 45 percent. In reading the percentage of eighth grade students scoring "advanced" or "proficient" increased from 14 percent to 43 percent, but the percentage required to achieve AYP was 54 percent.[11] Because Baker had not achieved AYP in any year, the school faced possible sanctions including restructuring, requiring all faculty and/or administrators to resign, or reconstitution as a charter school.

One of the laudable goals of NCLB is to make student achievement visible. To this end, NCLB mandates that schools disaggregate performance according to student demographics. In order to make AYP, schools must meet proficiency requirements for subgroups, including disaggregating student performance by economic status, race/ethnicity, status as English Language Learners and students with special needs. Many educational researchers and policy analysts have argued, however, that it is unrealistic for *all* students to reach proficiency by 2013-14. Schools with large numbers of minority and low-income students, students with special needs, and English Language Learners—all groups that disproportionately attend urban schools—are especially disadvantaged.[12]

---

[10] Baker also failed to show sufficient growth according to Safe Harbor or the Pennsylvania Performance Index.

[11] PSSA Scores. School Profile: Baker Elementary School. School District of Philadelphia, 2007.

[12] Hess, F. & Finn, C. (2007). Conclusion: Can this law be fixed? A hard look at the NCLB remedies. In F. Hess and C. Finn (Eds.), *No remedy left behind: Lessons from the first half-decade of NCLB.* Washington, DC: AEI Press; Linn, 2003; Novak, J. & Fuller, B. (2003). Penalizing diverse schools? Similar schools, but different students bring federal sanctions. *Policy Brief.* Berkeley, CA: University of California; Rothstein, R., Jacobsen, R. & Wilder, T. (2006), Proficiency for all: An oxymoron. *Education Week, 26*(13), 32, 44; Sunderman, G. (2008). Introduction. *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; The School District of Philadelphia's Accountability Review Council has reported that all district schools are not on course to meet the 2014 proficiency requirements.

Research conducted in 2006 on the characteristics of Philadelphia district schools that achieved or failed to achieve AYP demonstrated that the following four characteristics corresponded to a decreased likelihood of making AYP.[13]

1. Percentage of low-income students: Forty-six percent of district schools with lower than average poverty rates made AYP. Nineteen percent of schools with above average poverty rates made AYP.

2. Number of subgroups: Schools with two or fewer subgroups were more likely to make AYP than those with more subgroups. Seventy-eight percent of schools that made AYP had two or fewer subgroups. Only 34 percent of schools that had three or more subgroups made AYP. Schools that had special education or English Language Learner subgroups were even less likely to make AYP.[14]

3. Grade configuration of schools: Elementary schools in any configuration (K-4, K-5 or K-6) were more likely than K-8, middle, or high schools to make AYP. Forty-five percent of elementary schools made AYP, compared with 21 percent of schools with higher grades.

4. School Size: Schools with an enrollment of fewer than 700 students were more likely to make AYP than schools with a larger number of students. Eighty-one percent of schools with 700 or fewer students made AYP. Forty-one percent of schools with more than 700 students made AYP.

Baker had a greater percentage of low income students than the district average, had more than two subgroups, was a K-8 school, and had close to 700 students. Therefore, it was more likely than not that Baker would be among the district schools that had not made AYP.

Case Study Parameters

This case study is based on data that I gathered at Baker as a part of the team, led by Research for Action (RFA), who conducted the "Learning from Philadelphia's School Reform" study that examined various district changes under the state takeover between 2002 and 2007. NCLB and the state takeover occurred almost simultaneously. Given NCLB's and the district's emphasis on "data driven instruction," one of the topics RFA examined in this study was how schools learn to use student data to improve instruction.

---

[13] Travers, E. & Blanc, S. (Jan. 8, 2007). School yardstick fails to measure up. Philadelphia, PA: *Philadelphia Inquirer,* B, 2.
[14] This finding corroborates findings by Balfanz, R. & Legters, N. (2008) NCLB and reforming the nation's lowest performing schools. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press and Linn, 2008.

Baker was one of ten district schools that the team examined in depth in our effort to learn about the topic. Between 2005 and 2007, I conducted 22 interviews at Baker—repeated interviews with the principal and key members of the school's Leadership Team, as well as individual interviews with 8 teachers, three of whom were also on Baker's Leadership Team. In addition, I observed a total of 11 grade group meetings, leadership team meetings and half- and full-day professional development sessions. I also gathered information from written materials such as the School Improvement Plan and protocols for lesson planning and data analysis. These sources enabled me to get a rich picture of Baker's shared educational goals and values. They also provided valuable information about instructional decisions, assessment practices, data use, and supports in place for teacher and student learning.

As the research progressed, I was increasingly impressed with many aspects of the school's practice, and I often went beyond the research team's formal protocols about data use and probed perspectives on instruction and learning more generally. In many cases, information from several sources—principal, Leadership Team members and teachers—was strikingly consistent, suggesting that what I was learning about the school came from shared perceptions and experiences. Because the research design did not include classroom observations, I was not able to verify whether classroom practice reflected what I learned from these other sources.

At several points, I discussed my observations and interviews at Baker with other members of the research team. Team discussions that focused on cross-case analysis contributed to this case study. The case also benefits from and builds on extensive research that RFA conducted at Baker prior to the takeover, and illustrates the importance of analyzing current school practices within a historical perspective.[15]

School History

When the state took over the School District of Philadelphia in 2002, Baker was on the list of 86 low-performing elementary and middle grade schools slated to be paired with a total of seven outside providers[16] and required to adopt their provider's model. However, because Baker's test scores and those of 15 other schools showed statistically significant improvement between 1998-99 and 2001, they were removed from the list. From then on, this group of schools was informally called the "sweet sixteen."

The "sweet sixteen" schools, like the low-performing schools that were paired with an outside provider, received additional per pupil funding. However, these schools did not have to follow a specific reform model. They remained under district supervision and were required to use the district's Managed Instruction System, including a Core Curriculum and pacing schedule, Benchmark assessments, and SchoolNet to access data.[17] At the same time, these schools were able to make budget decisions using their

---

[15] Christman, J. (Spring, 2002). *Baker middle school.* Philadelphia, PA: Consortium for Policy Research in Education & Research for Action.

[16] These included Educational Management Organizations (EMOs), both for-profit and non-profit, and university providers.

[17] Christman, J., Neild, R., Bulkley, K., Blanc, S., Liu, R., Mitchell, C. & Travers, E. (2009). *Making the most of interim assessment data: Lessons from Philadelphia.* Philadelphia, PA: Research for Action.

extra funds, such as adding and reconfiguring staff, and developing special programs and interventions to address student needs.

When the state took over, Baker was a middle grade school. In 2005-06, the district decided to move away from the separate elementary and middle schools model toward a combined K-8 school model. The district converted many middle grade schools to K-8, including Baker. Baker and a nearby elementary school that was also becoming a K-8 school traded half of their students. However, the existing faculty wanted to remain at Baker. Thus, a group of middle grade teachers who had elementary certification volunteered to teach in the elementary grades.

The transition from a middle to a K-8 school was disruptive, but the new elementary teachers were committed to staying at the school. The district did not provide professional development or training about the K-3 curriculum, pedagogical strategies for teaching in the elementary grades or ways to address the developmental and behavioral needs of young students. Thus, at the outset, these teachers had to learn by doing and share what was working with each other. In the middle of the fall semester, a Reading First Coach was assigned to Baker. She was a "godsend" for K-3 teachers and provided critical support and resources, especially in the area of reading/literacy.

Before NCLB and the state takeover of the district, Baker had been involved with two comprehensive school reform groups: Johns Hopkins University's Talent Development and the Coalition of Essential Schools. The work of these two comprehensive reform groups at Baker had complemented each other.

Talent Development introduced a school-wide curriculum in math and reading. At the time there was no district-wide curriculum, so this offered the school a coherent academic program, which was a significant contribution, and helped teachers be on the same page about what should be taught. Talent Development also provided intensive professional development for teachers, and content coaches worked with teachers in their classrooms.[18] The Coalition of Essential Schools focused on leadership development, providing in-depth professional development for the school's principal and Leadership Team. It also encouraged teachers to use instructional strategies that promoted critical thinking and to evaluate learning through a variety of performance-based assessments such as writing, projects, and presentations.[19]

> **The principal and members of the Leadership Team were unwilling to make achieving AYP their main goal: to focus their instruction around preparation for the high-stakes state test, as many failing schools have done.**

Talent Development and the Coalition of Essential Schools had influenced the way Baker administrators and faculty thought about curricular issues, teaching strategies, types of assessments, and ways to improve learning. Building on the work of these reform

---

[18] Herlihy & Kemple, 2004 discuss the elements of the Talent Development model. Mac Iver, D. & Balfanz, R. (2000). Transforming high-poverty urban middle schools into strong learning institutions: Lessons from the first five years of the Talent Development middle school. *Journal of Education for Students Placed at Risk, 5,* 1- 2. Baker was one of the Talent Development schools and showed significant improvement in student achievement as a result of Talent Development's work at the school.

[19] For examples of these and other best practices used in the Coalition model, see "CES Resources," at www.essentialschools.org

projects and the leadership of the former principal, the current principal and other members of the Leadership Team were unwilling to make achieving AYP their main goal—to focus their instruction around preparation for the high-stakes state test as many failing schools have done. They were determined not to abandon their philosophy that education must foster life-long learning, student inquiry, application of concepts to the real world, and a focus on the whole child.

Meeting NCLB Requirements AND Remaining Faithful to the School's Educational Philosophy and Practices

Pennsylvania uses PSSA results to hold schools accountable for meeting NCLB requirements, and the district holds principals accountable for PSSA achievement as part of the Principal Review process. In addition to this accountability role, the district believes that student performance on the PSSA is valuable for deciding instructional goals at the school and grade level as part of creating the School Improvement Plan. As in most statewide assessments, however, results come after the end of the school year in which the test is administered. Thus, teachers cannot use the results to guide instruction in the current year.

In order to provide data to teachers that are more actionable, the district uses Benchmark tests, aligned with the Core Curriculum, that address the Pennsylvania standards. Benchmark tests are given every six weeks to monitor students' progress as well as to indicate areas of weakness at the grade, class, and individual student levels. The expectation is that during the 6th week of each cycle, benchmark results will provide data to inform re-teaching of topics in the Core Curriculum.

Improving the performance of *every* student in the school was key.

Baker's leadership believed that achieving AYP and addressing district curricular and assessment requirements with fidelity was essential. They used the Core Curriculum and Benchmark process to improve performance on the PSSA. However, the need to make AYP was not an end in itself. Making AYP was rarely discussed by Baker teachers as a factor that motivated their teaching. The school's goals included helping students apply what they learned in classes to the outside world and preparing all students for life-long learning rather than short-term achievement on the PSSA. Improving the performance of *every* student in the school was key.

Baker's principal had experience with a variety of assessments which embedded evaluation in instruction rather than simply using it as a tool to make summative judgments. His experiences with, and belief in, this form of evaluation led Baker to use a variety of assessments that provided more nuanced and timely information about individual student learning. These included samples of student work, writing, projects, and oral presentations.

The principal described the challenge of using two approaches to measure student performance—one that emphasized outcomes on a standardized state test for purposes of accountability and the other that used data from a variety of performance assessments in order to improve student learning.

Now, this is what I've learned. The problem is at this time I think you have to have a happy balance. Because the data that we have to attach ourselves to—that relates to students' standardized performance—is important to the teachers because for some of our teachers it has helped us to understand that, as good as we think we are, there's a lot of room for improvement…. On the other side, however, is that other part of the student that you can't leave out of the mix.

The one thing that I tell you, which is absolutely necessary, absolutely necessary, is that there's got to be a greater emphasis in looking at student work. And when I say student work, or let me say, a body of student performance, which goes beyond just simply the standardized test reports. In order to understand what you might need to do to help the kid to be more successful, you've got to have other data resources.

The School Growth Coach, assigned by the district in 2006-07 to support Baker's efforts to make AYP because it was in Corrective Action II, commented,

The pressures that I feel. Well, AYP. I really want to—I would like to see this school make AYP, and for no other reason that it is in a fourth year Corrective Action. And then after that, [after a school doesn't make AYP] I think they're going to get stricter. They really can reconstitute your school. And I think it would be a travesty. Because the numbers are not indicating what goes on here. It's really not…And for one to even consider reconstituting a school like this, or to even be in Corrective Action because of the numbers, instead of asking yourself, 'Okay, the school has a good culture. The teachers seem to be working hard. Why hasn't the school made it? What supports can I put in place to make—perhaps to bring closer alignment with the goals of AYP with what the school is doing?'

The principal and members of the Leadership Team were distressed, as was I, that Baker was labeled a "failing school" according to performance on a single standardized state test, as required by NCLB, when it was implementing what many educators considered "best" educational practices.

The Establishment of a Learning Community for All

Based on my field research at Baker, I became more and more convinced that it was an exemplary urban school—that it was a school with robust capacity and a learning community for students as well as adults. Below, I discuss the factors that contributed to my very positive impressions of the school.

I became more and more convinced that Baker was an exemplary urban school— that it was a school with robust capacity and a learning community for students as well as adults.

School Leadership and Teacher Learning

At Baker, the principal and other members of the Leadership Team were a powerful presence in the school and made the school a learning community that deliberately supported the growth of adults, which in turn supported student learning. Several of the positive leadership practices that existed at the school had roots in Baker leadership prior to the state takeover.

In 2006-07 the Leadership Team included the principal, the assistant principal, the Math and Literacy Content Leaders, the Reading First Coach, the School Growth Coach, the Roster Chair/Testing Coordinator, the three Instructional Climate Leaders (multi-year grade group leaders), and the physical education teacher, who doubled as the informal school disciplinarian. The school nurse and the guidance counselor also attended occasionally. Most members of the Leadership Team had several years of experience at Baker.

The principal had been at the school for more than 20 years, beginning as an art teacher, This was reflected in the school's unusual emphasis on the arts—due to district budget constraints, many schools had eliminated art programs in recent years. In turn he had become a dean, then the assistant principal, and finally the principal the year before the state takeover. When he was the assistant principal, he was mentored by his principal, who led the school during the implementation of Talent Development and Coalition of Essential Schools reforms, programs which continued to influence Baker after she left. The principal acknowledged that the former principal was critical in his development as a school leader and greatly appreciated her mentorship.

In his interactions with me, the principal was extremely candid, reflected critically on his educational philosophy and his experiences since NCLB, and said he welcomed my presence at Baker. When I called him to set up interviews, he often joked that "no time was a good time," because he was so busy. Once the often-lengthy interviews began, he gave his full attention and in numerous instances raised related issues in response to the questions I asked. At the end of one interview he said it had made him think "so hard." On another occasion, even though the interview had run over the time he anticipated, he said that he valued the opportunity to talk with someone who knew about teaching and learning.

The principal was central in setting the tone at Baker. He modeled the kind of respectful and caring interactions he expected from faculty and staff. He was perceived as a skillful, supportive administrator who had a complex understanding of teaching and learning. Yet, he acknowledged that he did not "operate under the guise that I know everything" and said he relied on the input of others. For instance, he arranged to meet informally with

other principals he considered "critical friends" to discuss school issues and problems. He spoke about making change happen.

> *You have to admit that in some cases, unless you develop the vision for the product, the product never happens. Now sometimes the vision is one thing, and then the details of how the product is going to turn out is something else. So you make sure that you utilize key people around you that know how to put the structure in place.*

The Leadership Team met weekly for almost an hour before school, in what was called "the huddle meeting" and at other times on an as-needed basis. The agenda of the meetings ranged from discussion of Benchmark data and planning educational interventions to discussion of school policies, school climate, testing procedures and planning professional development sessions. In meetings, different members of the Leadership Team facilitated parts of the agenda. Issues were seriously considered and in some cases debated. Team members also often complimented each other on their suggestions. Collaboration and collegiality characterized the interactions of team members.

The principal encouraged shared leadership among team members. In meetings, there was clear give and take, with information, questions, and ideas from the principal and other members of the Leadership Team going back and forth. In one meeting I attended, for example, the principal had not been in the school the previous day, so the Reading First Teacher volunteered to collect topics to put on the "huddle" agenda. The meeting covered a wide range of topics and was facilitated by many team members.

The principal encouraged shared leadership among team members.

Agenda
- Procedures for Thursday and Friday Test Prep (Literacy Leader/Testing Coordinator)
- Implications of district budget cuts for Baker (Principal and Ass't. Principal)
- Information about a new state directive for teachers to focus on "eligible content" (School Growth Teacher)
- Flu shots (School Nurse)
- Identifying students for CSAP (Guidance Counselor)
- New teacher meetings (School Growth Teacher)
- Upcoming Benchmark Tests (Literacy Leader/Testing Coordinator)
- Incentives for Attendance (Phys. Ed. Teacher, who was also the informal school disciplinarian)

An edited version of the "huddle meeting" that dealt with incentives for attendance reflects the positive and constructive nature of these meetings.

> Phys. Ed Teacher/Informal School Disciplinarian: *Perfect Attendance initiative starts today. The Baker cash is in.* [Baker cash can be redeemed for prizes twice a year.]

> Principal: *Thanks for making sure that these students [with good attendance] are remembered….ICL's [Instructional Community Leaders] need to remind folks that, if their classes meet attendance goals, they get to see a movie per month. If the whole school meets the*

goal, they get to have "Dress Down" Friday. [Students are required to wear uniforms on regular school days.]

Are there any problems with it? The kids value it, but I know some adults have problems. It is an important incentive for kids. But we need to keep checking.

Assistant Principal: *Dress down shouldn't mean behavior is down…. I saw some lax behavior. I saw some hats, sunglasses, etc. Dress Down Days are not lax days. They are regular instructional days. No eating in class.*

Principal: *You have a good point. Just want to keep my finger on your pulses. Staff have to understand that dress down day isn't take-off day either.*

ICL (Grade Group) Leader, K-3: *Yeah, the kids' behavior IS lax sometimes.*

School Growth Teacher: *I have an idea for a "caught being good" initiative [for Baker]. An example is Room 207 who was in a beautiful line. I took a picture. If you see an example, take a picture. …*

ICL (Grade Group) Leader, K-3: *That is a great idea. We can have a bulletin board on every floor, so kids can see them.*

Principal: *I think you can be "caught being good" all over the school. Caught being good can be inside or outside of classes and the school building. We should have banners promoting caught being good initiative. That's what (ideas) you get when you have new people outside of the Baker culture….* [School Growth Teacher had only been at Baker since the beginning of the school year.]

Assistant Principal: *What is the policy on black (plastic) bags?*

Principal: *Kids can just keep sandwiches, not junk. No candy, gum, bottles in school. I know that parents say they spent good money, but we need to take junk food away from kids…. In some classes junk needs to be swept up. It is extra work for the custodians.*

[Bell rings for classes to begin.]

Principal*: OK, group. Thanks.*

The principal and leadership team worked collaboratively and everyone took responsibility for making the school a learning community for faculty. Members of the Leadership Team regularly worked with teachers: examples included observing classes and giving feedback, modeling instructional strategies, problem solving, discussing data and the implications for differentiated instruction, and providing opportunities to reflect on practice. Teacher leaders provided non-threatening support. The teachers all gave positive examples of their interactions with at least one, and often more, of the teacher leaders.

Teachers felt comfortable sharing their questions and asking for help from each other and members of the Leadership Team, especially the Content Leaders, Reading First Teacher, and School Growth Coach. They said that they also received feedback from the principal,

who encouraged them to learn from other members of the staff. A teacher commented about the principal's support:

> *So, when we are doing something on constructed responses, doing test prep [in seminar], or something, the principal strolls around and, you know, and makes sure to maybe highlight the classroom that he walked into and say, 'If you need to see what's going on in seminar [a daily thirty minute at the beginning of the day, when students were divided into two groups per grade for differentiated instruction], go to so and so's room because, you know, this is a person who's doing something positive in seminar'…So that's why he'll highlight and say, ''If you need to, you need to go visit so and so for some advice or something.'*

But the principal realized that such suggestions were not necessarily sufficient. He explained that some teachers were more motivated than others to try out new approaches, such as using SchoolNet to retrieve and analyze data reports.

> *Ideally, I would like to say, as an adult, a professional adult, I just need to give you the charge. But it doesn't work like that. You've got to give them a framework that puts them in a position that requires them to give you feedback. And then you are most likely going to get the feedback. So, we've created templates where they're required to give us information [about classroom plans] where they're also required to connect [plans] with the data.*

Walkthroughs, observations of teaching practice, and direct feedback by the principal and Leadership Team members helped teachers to reflect on and improve their teaching. The principal explained,

> *We've found with doing our informal walkthroughs that there's just too much teacher-directed instruction, too much teacher talk. And you talk to the teachers one-on-one or in groups, say in their ICL meetings, 'How are you doing this?' They really believe that they're giving the kids the opportunity [to do higher order thinking]. But when we give them examples, 'Did you know that during that class when you thought that you were really encouraging kids to be engaged that you were really responding to just four kids?' You've got to understand that if you're only getting four hands, what's happening to the rest of them? If I'm asking a kid questions that are so close-ended that they're giving me just simply facts, and they're not talking to me about how I came to that answer, whether the answer was right or wrong, you're not really accomplishing a large part of your goal.*

The School Growth Coach quickly realized that the leadership of the school was exceptional.

> *When I think of the kind of leadership that the principal offers, the Leadership Team has, if you were to write a leadership text, it couldn't be better than coming here to observe. This is a case study of what leadership should be like. It really is.*

Faculty Learning Together

Faculty contributed significantly to making Baker a learning community. The faculty was distinctive in terms of its longevity and attachment to the school. Teacher turnover at the school had been minimal since the takeover. The lengthy tenure of many faculty provided important continuity and the possibility for the development of common goals and efforts. Teachers praised the cohesiveness of the staff and the willingness to help one another, qualities that I learned about and admired.

Teachers worked together with grade partner pairs to plan lessons and share teaching strategies that they found helpful for meeting student needs. Informal meetings between grade partners provided the opportunity for teachers to discuss data and instructional strategies for the grade as a whole as well as for individual students. As the Literacy Leader, who also taught 8th grade, explained,

The lengthy tenure of many faculty provided important continuity and the possibility for the development of common goals and efforts.

> *We [she and the other 8th grade literacy teacher] share writings that we think are great. And I will go to X and say, "I think this essay is great. Evaluate it for me." And we do that all the time, because we're using the same curriculum. We plan our lessons together…So we work together all the time. There are a lot of [grade pair] teams that do that.*

Teachers reflected on their teaching in structured ways. They met in purposive groups in professional development sessions to examine data and consider implications for classroom instruction. Classroom teaching was not only an individual endeavor. Rather, it was an activity that faculty members could observe and learn from. Teachers were encouraged to observe "best practice teachers" in their grades and/or subject area. In this sense, teachers deprivatized their instruction.

The public, collaborative stance toward improving instruction was reinforced by thoughtfully planned professional development. Each year the district scheduled several professional development sessions for all schools. While the district typically forwarded generic scripts for these sessions, Baker regularly adapted the script or developed their own plans, because they believed that professional development was most useful when it addressed particulars of their school. The principal occasionally distributed articles about subjects such as how culture affects learning, as "reflection pieces for teachers to think about what it is they could do to make [learning] more meaningful."

Many professional development sessions focused on analysis of Baker data from the PSSA and Benchmarks and their implications for instruction. A teacher explained,

> *Well, this year, like we've looked at this data in and out, backwards, upside down, the Benchmarks, the test scores, everything. So, a lot of our staff development has been on this data. So, we can see what kids need.*

In other sessions, professional development provided opportunities for teachers to learn by doing—to practice skills they could use in their classrooms. During professional development teachers worked in different small groups depending on the goals of the session: in grade groups (clusters of two-three grades), subject matter groups, and grade groups by subject matter. Members of the Leadership Team facilitated segments of each

session. At the end of most sessions, teachers were expected to complete a brief evaluation form.

For example, Baker designed its own professional development session on how to write and use higher order questions to improve students' critical thinking. The School Growth Coach discussed levels of higher order thinking skills and gave examples of higher order questions that "challenge" students to reason at higher levels. Teachers broke into grade groups and discussed ways to use higher order questions, wrote sample questions and from then on were expected to include higher order questions in their lesson plans.

In their classes, faculty regularly used PSSA rubrics to assess written work, both essays and responses to open-ended questions. Some professional development sessions focused on helping teachers to understand the rubric assessment process and how to evaluate student writing using rubrics. In small groups, teachers scored samples of writing according to rubrics, justified their scores, and attempted to come to a consensus for samples of writing.

A planning meeting for an upcoming professional development session that I observed in fall 2006 focused on how to deepen teachers' understanding of what various assessments indicated about student learning. For the session, teachers would be asked to bring:

- Benchmark data,
- Student work samples,
- Classroom assessments,
- Progress on writing constructed responses, and
- Reading assessments for younger students.

In the professional development session, teachers would be asked to compare the value and limit of various assessments to provide a full picture of a student's instructional needs. The planning meeting revealed the kind of critical thinking about assessment that I believe should occur in all schools. Many respected educators and educational researchers argue that using a variety of assessments, rather than a single high-stakes test, is essential to assess student performance in a reliable and equitable way.[20]

---

[20] Baker, E. (2007). *The End(s) of Testing.* Presidential Address. Chicago, IL: American Educational Research Association Meetings; Darling-Hammond, L. (2003, February 16). *Teachers College Record.* ID Number: 11109. Retrieved on February 8, 2009 from http://www.tcrecord.org; Darling-Hammond, L. (2008). Improving high schools and the role of NCLB. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; Hawley, W. (2008). NCLB and continuous improvement. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; Kornhaber, M. (2008). Beyond standardization in school accountability. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; Linn, 2008.

Use of Data to Improve Student Learning

Between 2005 and 2007, the purposeful use of student data from multiple sources (Benchmarks, writing rubrics, reading inventories, as well as a variety of classroom and performance-based assessments) was evident. Of these, the interim Benchmark assessments provided a common frame of reference. Teachers regularly analyzed Benchmark data in many ways and settings—during formal professional development and during informal meetings with grade partners, with Teacher leaders/coaches, and by themselves to make instructional decisions.

The purposeful use of student data from multiple sources (Benchmarks, writing rubrics, reading inventories, as well as a variety of classroom and performance-based assessments) was evident.

A teacher described how she used the district's Benchmark Item Analysis Report to make instructional decisions.

Which three standards did the kids do poorly on? And which kids were they?...Well, you know, when we're doing homework, when we're doing class work, the kids did pretty well on this particular standard. What was it that messed them up when they did the benchmarks, because we either do paper or computer. What was it when they read it that stumped them? Why did they get this standard wrong?

In order to use the Benchmark data to guide instruction not only at the classroom but also the individual student level, Baker had designed a tool that added value to the district's Benchmark Item Analysis form. It specifically asked teachers to look at individual students' Benchmark results after each six-week cycle and to indicate a plan of action for each of them, especially for low performing students. As the principal explained, "The teachers were told, "Do not give us a generic plan…. Know what data are attached to each student."

While teachers relied on interim Benchmark data to plan interventions for groups of students and individuals in their classes, most believed that they needed information from additional assessments to help them tailor day-to-day instruction more effectively. They, as well as school leaders, shared my concerns that responses to multiple-choice questions on interim benchmark tests did not reveal the breadth and depth of student learning and that such tests frequently are not adequate measures for students with different learning styles. Benchmark tests did not provide fine-grained diagnostic information that could be used continuously to guide planning.[21]

Information from a variety of additional assessments helped teachers to identify the ways individual students learned, to gauge students' academic strengths and weaknesses, and to design interventions for groups of students in a timely manner. Teachers used assessments that were embedded in regular instruction, including student writing, samples of student work, projects, and classroom observations to provide evidence about student learning that helped guide decisions about daily plans. In many cases,

---

[21] The debate about value of interim assessments, such as the Benchmark tests used in Philadelphia, to guide day to day instruction is discussed in Cech, S. (2008). Test industry split over 'formative' assessment. *Education Week, 28*(4), 1,15; Baker, 2007 critiques the claim that interim assessments/benchmarks are formative assessments.

performance assessments were seamless parts of lessons. One teacher gave an example of how listening to student discussion in small groups helped her assess students' understanding.

Lesson plan templates included questions that prompted teachers to be aware of individual student needs as they designed lessons. As a teacher described,

> *Well administration, with the way they set up the lesson plans, kind of geared us in the direction of addressing kids based on skills, not necessarily based on levels, because two kids could be only "basic" but have different needs. They really pushed for us to address kids as individuals and their individual needs. And when they did that and presented the lesson plan in that way, we had to write specifically which children we were helping with which big skill.*

The school encouraged student reflection about and assessment of their written work and Benchmark data. For instance, the leadership team had developed a "Student Voice" form that asked students to examine their Benchmark data and consider the implications for subsequent learning. Some teachers also asked students to explain their reasoning on test items they answered incorrectly. A teacher reported,

> *Like when I work with the kids, [lowest group] I'll go back and I'll say, well, "John, you chose A? Why did you choose A? What made you think that it was A? Go back to the story and show me why your mind was choosing….Because they give you, you know, what they chose. That's extremely helpful.*

The principal stressed the importance of students understanding the implications of their own data.

> *If you're going to help the kids see any significance in why you're doing some of what you are doing, you've got to make sure that they understand what their own data suggests, whether it's PSSA, benchmarks, their classroom achievement data, whatever it might be….And then make sure that they understand how they can make improvement, because that was the whole point behind having the teachers keep portfolios for the kids and check them at intervals during each marking period.*

School-wide Interventions to Increase Student Learning

Baker used PSSA data to identify state standards on which student performance was weak in the school as a whole as well as in individual grades. On both the multiple choice and constructed response portions of the PSSA, Baker students did poorly in both reading and math.

Benchmark tests, aligned with the Core Curriculum and the PA standards, were intended to simulate questions on the multiple-choice section of the PSSA. The Benchmark tests, however, did not include any PSSA constructed response-type questions that required students to explain their answers. Given their belief that meaningful instruction should foster skills for life-long learning, the principal and members of the Leadership Team decided a key instructional goal should be improvement of students' ability to write and

to answer constructed response, open-ended questions, rather than simply preparing students to do well on the multiple choice section of PSSA.

The school implemented two innovative school-wide instructional strategies to address weaknesses in writing and reasoning.

The school implemented two innovative school-wide instructional interventions to address weaknesses in writing and reasoning. These were "Thursday-Friday Test Prep" and "Large Group Seminar." Both interventions were responses to weak performance on the PSSA but both were intended to support meaningful learning goals rather than test prep as an end in itself. In these interventions, test prep was embedded in authentic classroom instruction to develop writing, problem solving and higher order thinking skills. It was an integral part of lessons rather than narrowly focused on skills needed to do well on the test.[22] A teacher described the process:

> *For example, students read a story problem and they have to explain how they got the answer. A lot of our kids can get the answer, but they can't explain how they got it. So that's become one of our main focuses that comes from the data...Writing, writing, writing is like our big thing....They want to see writing in every single classroom, every single day.*

Thursday-Friday Test Prep

On the constructed response portion of the PSSA, in which students wrote answers to open-ended questions, a sizable portion of Baker students received a score of zero or one out of a possible three points in 2004-05. To address students' weaknesses on constructed response items, Thursday-Friday Test Prep was instituted in the middle of the 2005-06 school year. (The name of the intervention is unfortunate because it suggests the common meaning of "test prep," where a school focuses on narrow skill-based instruction, when the purpose of Baker's Thursday-Friday test prep was to improve writing and thinking skills that would be useful in the future.) This decision was significant because while the constructed response portion of the PSSA contributed a relatively small percentage of a student's overall PSSA score, Baker leadership decided to adjust the weekly schedule to implement this curricular intervention.

Each Thursday, for a half of their literacy or math periods, students answered PSSA-type constructed response questions under test-taking conditions. On Fridays, teachers returned students' written responses and students scored samples according to the PSSA rubrics. Teachers used a variety of strategies to teach students how to assess writing. A teacher might put samples of student work on an overhead that included common problems and ask students to score them and make corrections. She might direct students

---

[22] Black, P. & Wiliam. D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy and Practice, 5*(1), 7-74; Chapius, S. & Stiggins, R. (2002). Classroom assessment for learning. *Educational Leadership*, *60*(1), 40-43; Stiggins, 2001; Wiggins, 1998; Wiliam, D. (2001). An overview of the relationship between assessment and the curriculum. In D. Scott (Ed.), *Curriculum and assessment.* Greenwich, CT: JAI Press; Wiliam. D. (June, 2004) *Keeping on track: Integrating assessment with instruction.* Invited Address to the 30th Annual Conference of the International Association for Educational Assessment, Philadelphia, PA. Authors discuss how types of performance-based assessments can be integrated into curriculum and instruction and lead to improved learning outcomes.

to evaluate sample responses in small groups and discuss reasons for the scores they gave. Students also self-evaluated their own answers.

The principal discussed the plan shortly after it was implemented,

> *The purpose of it is to give kids opportunities to look at their work against rubrics, because a lot of times, what we're finding is that our kids have a base of knowledge, but their problem is communicating the knowledge....After teachers look at student work for their own assessment. What is it kids seem to be getting? What is it that they're not getting?' They give samples of written responses to students and ask them, 'What makes a piece of work a 1 or a 2 versus a 3 or a 4?' (There were three levels of rubrics in reading and four in math)*

According to the Literacy Leader, students took the process seriously.

> *You teach [rubric scoring] to the whole class, and they rate each other's [work] much more stringently than we would... 'Does it have a closing sentence? Are there three pieces of information that you can find in the text that you have cited? Have you answered the question properly?' Give the answer to another group, and see how the other group rated it. And if it's not the same, then the kids fight it out.*

In spring 2006, Thursday-Friday Test Prep, instituted two months before the PSSA's, was a work in progress and not expected to significantly affect Baker's constructed response scores on the state test. In spring 2007, the impact of the Thursday-Friday Test Prep plan, however, was clearly evident. For instance, from 2006 to 2007, scores on open-ended questions in reading improved significantly in every grade, except fifth. Between 2006 and 2007, the percentage of students receiving scores of 2 or 3 in grades four, six, seven, and eight increased between 11 percent and 25 percent.

Improvement in constructed response scores confirmed my belief that providing an innovative writing sequence that enabled students to reflect on their own writing and score students' writing samples was a powerful learning strategy. It engaged students as agents in their own learning, which is a hallmark of student-centered and constructivist practice. Professional development, as discussed earlier, provided teacher training in the use of rubrics. In combination, the Thursday-Friday Test Prep plan and professional development addressed the writing skills of all students, no matter what their initial levels of writing.

Large Group Seminar

For the Large Group Seminar that began two months prior to the 2007 PSSA's, teachers identified students in grades 3-8 who were "advanced," "proficient," and those they believed could become "proficient" for special instruction, a strategy designed to enable higher percentages of students to achieve proficiency. These students received enriched instruction in reading, writing, math and higher order thinking. Seminars composed of students from two classes met for an hour and a half. Each seminar included a "warm-up activity," a "thinking out of the box" activity and stressed how what students learned could affect their lives. These seminars also included learning about the PSSA and strategies for answering multiple choice and open-ended questions.

The School Growth Coach, Content Leaders and "best practice" teachers taught the seminars. Together they planned what sounded like exemplary lessons. After each seminar, they critiqued the lessons and planned for the next session, strategies I would argue are powerful teacher learning approaches. The principal explained they used seminars taught by "best practice" teachers to model higher order thinking. He explained, "Our kids will only do what we give them the opportunity to do. And if we're not engaging them on higher levels, they're not going to respond on a higher level." The School Growth Coach said the seminars were "a big deal" and students were upset if they had to miss any.

In the two months prior to the administration of the 2007 PSSA, teachers also were given a special lesson plan template to use. This template asked teachers to think critically about their instructional planning. The lesson plan template included questions such as, "How are you going to accomplish your goals?" "What activities are you using?" "What depth of thinking are you challenging children to have?" "How are you giving them opportunities to express themselves, either orally, or through projects, or through writing?" Teachers were expected to focus on "eligible content," designated by the state, for improved PSSA performance. On Fridays, during these two months, all students also participated in a Project Period, based on a topic the class had determined during the week, a practice that encouraged students to set their own tasks and take responsibility for the projects' outcomes.

The principal explained this interdisciplinary, collaborative approach.

> *The project approach—it's a process that they can use over and over again as they are progressing through high school, if they're going to go to college, whatever it is. The problem that we're finding in trying to do things the way the school district has outlined it, with giving them just simply a tool kit of strategies, is that the kids can have the tool kits. But unless they understand the real-world application of it, then it means nothing…So in the end, they are not just being given a strategy, They've been given a life-long tool.*

Although the instructional interventions were developed in response to the school's poor performance on the PSSA, they went well beyond common test prep, which is typically decontextualized.

Although the instructional interventions described above were developed in response to the school's poor performance on the PSSA, they went well beyond common test prep, which is typically decontextualized. At Baker, making AYP was not an end in itself.[23] It is possible, however, that this approach contributed to the fact that Baker had never achieved AYP according to the status achievement model because it did not regularly focus on short-term student outcomes.

---

[23] Baker's identification of these students for enriched instruction resembles in some respects the practices of identifying "bubble kids" who could become proficient. However, the seminar's rich content and innovative teaching strategies did not resemble the narrowly focused skill-based preparation of such students that characterizes many schools where a major goal is to increase the percentage of proficient students.

School Leaders Reflect on the Role of NCLB
in School Improvement

NCLB directly and indirectly had important pedagogical and psychological consequence at Baker. Reactions to the broad goals underlying NCLB were generally positive, and some of the initiatives that grew out of NCLB were also seen as supporting school improvement.

In order to improve student performance and hold schools accountable, NCLB mandates that all teachers be 'highly qualified." It encourages states and districts to use data to make pedagogical decisions, and to provide supports for low performing schools. The Literacy Leader praised the NCLB requirement for hiring highly qualified teachers.

Reactions to the broad goals underlying NCLB were generally positive.

> *To be at all honest, NCLB is making the teachers more qualified. You're not going to have a teacher in an eighth grade reading setting that is not trained to be there. So, I mean, I think it's fine. Now the students are also being asked to be accountable.*

The principal and Teacher Leaders in Math and Literacy agreed that the district's response to NCLB had spurred an increase in the amount and quality of data use at the school. They explained that Baker had used student data to make instructional decisions well before NCLB, when the school worked with outside organizations on school reform projects. They reported that since NCLB, however, the district's emphasis on analysis and use of student data to make instructional decisions had definitely increased the amount of time that they, and Baker teachers, looked at, and used, data. As the Math Teacher Leader explained,

> *We always had a leadership team, and we spent time looking at data. The former principal was very big on looking at data to improve instruction and student achievement. After she left, our current principal, who was the assistant principal under her, continued in her footsteps. She mentored him, and he also believes that focusing on data is central to improving student learning. I think as a school we have always seen data as part of what we are all about, but in the past couple of years, with the PSSA, AYP and the Benchmarks, we have been even more focused on data. As a leadership team, we discuss results of student performance in all kinds of ways.*

Professional development at Baker regularly focused on ways to analyze and use student data from both standardized tests and performance assessments. As the Literacy Leader reported in 2006,

> *We analyzed Benchmark data in all kinds of ways to examine student progress, and to access where students were weak and strong. We also looked at samples of student work to see how these compared to performance on the benchmarks. I believe that looking at student work, in addition to the Benchmark data, is essential for teachers to gain information about a student that can't be measured by Benchmarks. We*

use professional development days and special after school meetings to focus on what can be learned from data and how to use it to work in classrooms. But teachers are not confident about how to evaluate student work—to give a piece of writing a 3, 4, etc. So, we decided to put together a handbook for next year with samples of student writing at all levels of proficiency by grade, so teachers would have some criteria and guidance to say, this piece of work is a 1, 2, 3, 4. I think teachers like the idea of having a handbook.

The principal explained that using standardized test results to make instructional decisions was challenging to teachers. He noted that the availability of data from the district's data management system, School Net, was only the initial step. In order to dig deeply into the data, he was adamant that teachers needed additional guidance and tools regarding how to utilize it. In an interview in 2005 he showed me a template developed at Baker to ensure that teachers would both look at benchmark data and use it to guide instructional decisions.

The teachers have to be trained as to how to use data. The other issue is the apprehension that they have about using this data, because they feel inadequate about using it. What you have to do is create vehicles to force them to use the data. And that's the purpose of the template we've created....The data is available. "This is what you're required to do. And I know you're uncomfortable, but this is an indirect way of forcing you to read and use it."

As a result of NCLB and the district's efforts to support schools that were in Corrective Action, in 2006-07 the district assigned a School Growth Coach to work full time on-site at each of these schools. The role of the Growth Coach was to develop processes and pedagogical interventions to support the school's efforts to make AYP. As the principal discussed the Baker School Growth Coach in glowing terms and described the many ways in which he impacted the school, he echoed the views of almost every adult in the school.

Right now he is touching every part of the school in one way or another. He's heavily involved in the professional development as a leader, as an organizer, designer, whatever it might be. The man has tremendous strengths, and he knows how to communicate his knowledge.... The School Growth teachers, in essence, are people that were supposed to be attached to schools to support teachers, whether they were new teachers or old teachers—to be available to model for teachers, to pull together resources, so forth and so on. And here he does all of that plus. And I'm telling you, he's in classrooms working with kids and all of that. This is not the intent of the School Growth teacher. This is something that he's taken upon himself, because he knows that it's a natural outgrowth of how he can affect academic improvement in the school, and even instructional improvement.

Concerns about the Unintended Consequences of NCLB

While the principal respected the educational rationale for NCLB, at several points he spoke critically about the role politics played in its development.

> *No Child Left Behind isn't completely wrong, in the sense that at some point there needed to be something to hold American educators and American education to a level of accountability for its product. The problem is that right now No Child Left Behind is primarily driven by political forces. It's not driven by people who are really, really concerned about what's happening with American education.*

In many ways Baker leadership believed that NCLB and the district's focus on making AYP had negative consequences for the school and interfered with its attempts to meet the needs of all students. The principal captured the concerns of many at Baker who believed that NCLB and the requirement that all students be proficient by 2014 placed serious constraints on teaching and learning at the school. He felt that the requirement to make AYP had impeded important student learning and that, overall, NCLB had forced him to make decisions that he believed were not in the best interests of his students.

In many ways, Baker leadership believed that NCLB and the district's focus on making AYP had negative consequences for the school and interfered with its attempts to meet the needs of all students.

> *NCLB has changed in many respects a number of my approaches on how I would have run this school otherwise, and for which I think I would have better results, to be honest with you. Because you're not engaging kids through their interests and their talent. So, to answer that question, NCLB has radically changed or caused me to adjust my approach to how I would do this job if I had the freedom to do it otherwise.*

While the principal understood the rationale for the district's Core Curriculum and Benchmark tests and required his staff to fully implement them, he said that NCLB and the district's response to it interfered with rich curricula that were in place at Baker before NCLB. In 2007, he described an interdisciplinary project on the Harlem Renaissance that the second graders had just completed, culminating in a trip to the Harlem Museum because they were not a tested grade. He bemoaned the fact that under NCLB, project-based learning and field trips were significantly scaled back in order to concentrate on reading and math.

> *I mean the Core Curriculum has made an impact on just about everything we've done. In the past we've basically had a belief, a strong belief, in project-based learning. But because of this AYP and oversight and so forth, project-based learning has been put on the back burner. And unfortunately, we believe that that may have slowed our progress more than assisted.*

He was concerned that the district's stress on making AYP caused some teachers to be wary of shying away from the Core Curriculum and pacing schedule and that this reduced their use of more time-consuming, but richer, curricular approaches.

> *A number of the teachers, because they really didn't get a whole lot of professional development with the Core Curriculum—literacy and math two years ago, science, social studies this year—have more or less locked themselves in a nervous box. 'I've got to stay on pace.... I'm being judged on the basis of, the school's being judged, on the basis of whether we're doing the Core Curriculum.' And so, project-based learning has not been looked at the same way—that it could still be incorporated with the learning experiences of the children.*

He believed that the heavy diet of reading and math and after school remediation for low performing students, coupled with standardized assessments, undermined their engagement in learning.

> *You know, my concern right now is that [with AYP] we're alienating a large block of kids, too large a block to want to come to grips with. They' are looking at themselves as failing persons because they're not within a high-performance group based on these standardized tests.*

The principal feared that Baker's failure to achieve AYP year after year had affected teacher morale, even though he and other members of the school's leadership team did not emphasize sanctions that might be forthcoming if the school continued not to make AYP. At the end of the 2005-06 school year, he talked about possible teacher discouragement in the fall if the school failed to make AYP.

> *I think NCLB has affected teachers from a morale point of view. I think that if we didn't remain, as an administration, consistent across four years plus—positive about why we need to do these things—I think that we could have had a disaster here in student performance. Because a lot of the teachers just don't believe that the way NCLB is approaching educating kids is the only way we should be doing it.*

The Literacy Leader also expressed concerns in this regard.

> *I sure hope [we make AYP,] because at this point I don't know what else we can do to help our kids. We do 15 things that other schools haven't even thought of yet. And I sure hope that we've made a difference. But you can't ask—the teachers give 150 percent. And the kids seem to be giving it and their attendance is okay. And I just hope in August we find out that we did [make AYP].*

One of the key supports the school district provided for schools who had not made AYP was a School Assistance Team (SAT). The SAT was chaired by an individual appointed by the district and staffed by the school's principal and other members of the school's Leadership Team. The SAT process, including periodic formal observations of teachers, analysis of Benchmark data, and feedback to school faculty and the district, had the potential to significantly impact instruction and improve student learning. The principal had high hopes for the SAT process. On balance, however, he found the SAT process more harmful than helpful. The principal was frustrated because the district had not assigned "the appropriate individuals to assist" a process of productive school improvement process, and the SAT leaders assigned to Baker had played more of a monitoring than a supportive role.

In 2006-07, the principal talked frankly with the current SAT leader about the kind of feedback that would be helpful, and the situation improved.

> *The SAT team process is evolving in a process that it is becoming a little more meaningful for us because the team here is insistent on making it meaningful. The SAT review is becoming a process that's working for us because we have challenged the process.*

For the Principal, one of the most demoralizing aspects of NCLB's press for accountability and the fact that Baker had not made AYP for several years was his perception of how the district viewed Baker. He was discouraged that the district had not acknowledged the many excellent practices at Baker. At a professional development meeting in 2007, the principal spoke candidly about this reason to make AYP.

> *[We] need do everything in our power to achieve AYP. Whether you like it or not, or you believe there is too much emphasis on testing, we live in a world of accountability. Baker needs to demonstrate what we do to be taken seriously by the district. The district will only listen and validate what we do, if we do what we need to do this year. We ARE going to do what we need to do."*

In an interview that year, the principal reiterated this theme.

> *If you want people to appreciate the work you do, unfortunately, you've got to hit the mark. Because I believe the day this school hits the mark—if this school hits the mark this year, somebody's going to look a little more seriously at what we're doing.*

Complicated Choices

As suggested at various points in this paper, NCLB and the requirement to achieve 100 percent proficiency by 2014 according to the status achievement model presented serious challenges for Baker leadership. Based on their professional judgment, they responded by making a number of complicated choices about curriculum and assessment, as they struggled to maintain their commitment to progressive educational goals and practices.

Despite the fact that they implemented the district's Managed Instructional System—adhering to the prescriptive Core Curriculum, administering Benchmark tests in good faith, and using School Net resources to retrieve and analyze student data—school leaders had serious reservations about its potential as a vehicle, by itself, to support a meaningful teaching and learning environment. On the other hand, remaining faithful to the school's pedagogical principles proved to be extremely challenging in an era of high stakes accountability.

Any decision to divert focus from AYP came with a cost. The principal reported the tremendous personal and professional conflict he felt because the school had not achieved AYP. He was demoralized that the school

Based on their professional judgment, Baker leadership responded by making complicated choices… as they struggled to maintain their commitment to progressive educational goals and practices.

district would not acknowledge what he referred to as "best practices" at Baker until the school met AYP goals. He was adamant that concentrating on narrow, skill-based teaching, whose main purpose was to improve student achievement on the PSSA, was not a means for deep learning. He refused to use this strategy, although he was convinced that some successful district schools did exactly this to make AYP. At the same time, he worried about the sanctions the school would face if it did not make AYP, and the possibility that the district would transfer him from the school for which he felt tremendous responsibility and affection.

Nevertheless, in several instances, Baker's school leaders made conscious decisions to emphasize pedagogy for life-long learning, rather than pedagogy that was more narrowly focused on skills needed to do well on the PSSA in order to achieve AYP. For instance, the decision to implement project-based learning as part of the curriculum, albeit it to a limited extent, took time from the Core Curriculum but promoted student creativity, inquiry and collaborative learning. The principal lamented that, because of the Core Curriculum, the school could not do as much as they used to, but he was determined to offer students a few opportunities to engage in interdisciplinary projects that he believed contributed to students' development as life-long learners.

The Thursday-Friday Test Prep intervention was an example of a decision that was designed to support life-long student learning that disrupted the district's curriculum pacing schedule. Baker leadership changed the weekly schedule to implement this intervention because they felt it supported several of their key goals—the development of students' writing and higher order thinking skills and their ability to assess their own work—goals that would not have been met by more traditional test prep. Moreover, in the Thursday-Friday Test Prep they chose to focus their efforts on students' ability to write responses to open-ended questions, rather than on skills needed to answer multiple choice questions, despite the fact that the constructed response section of the test counted for a relatively small portion of a student's total score on the PSSA.

In terms of assessment, many Baker staff had reservations about the district's reliance on data about student learning from standardized Benchmark tests. They believed that these tests, administered every six weeks, could not show the full extent of what students with different learning styles and academic strengths and weaknesses knew and could do. Moreover, they believed that benchmark tests did not offer teachers more helpful, "real-time" feedback. Thus, in addition to the Benchmark tests, which were scored at the district, teachers often relied on a number of performance-based assessments that were more difficult and time-consuming to evaluate. However, such tests enabled them to design lesson plans that captured fuller information about the needs of individual students and groups of students in their classes.

Another difficult decision the principal and Leadership Team made was to candidly question the way their SAT leader conceived of his role, and to give him feedback about how he could be more useful—an approach fraught with complications because of his ultimate authority on the SAT. The principal, however, was determined to make the SAT process work for the school, and thus took the risk of asking him to give feedback that would be more helpful in the process of school improvement, and to play a more collaborative than monitoring role.

The school leadership was determined to serve the varied needs and learning styles of *all* Baker students. A number of practices described earlier are testament to their commitment to teaching and reaching every child. As the imperative to respond to the consequences of failing to make AYP became greater, the school leadership made some strategic decisions in an effort to meet AYP goals on the state's standardized test. Teachers were asked to identify the "chosen ten," (highest achieving students) in their classes and to provide some differentiated instruction for them. They provided more rigorous and enriched curricula to students who already were, or who they thought could become, "proficient" in order to meet short-term proficiency goals—a strategy used by many low-achieving schools who target "bubble kids" on the cusp of proficiency to meet AYP targets.

After the principal announced the plan for the Large Group Seminar at a professional development session in 2007, a teacher questioned him about the decision to privilege the instruction of higher performing students, saying it was challenging to work with "intensive" (lowest performing) students and that these students needed enriched instruction as well. The principal acknowledged the conflict he felt about allowing the need to make AYP affect his professional judgment and lamented that,

> *NCLB is creating a group of dysfunctional kids....But you can't move everybody up at the same time. We tried that approach—to support all students—but right now we have to think more strategically. We need to look at students who can move.*

His frank response that Baker needed to spend more of its human resources on students who were already "proficient" or thought to be moveable for purposes of making AYP revealed the challenge he faced. Although he admitted that Baker needed to make some strategic decisions in the short run, most of his public and private statements showed his commitment to focus on the life-long learning of *all* students.

In the final interview with the School Growth Coach, he discussed the dilemma of meeting the AYP performance requirements of NCLB, while trying to maintain the goal of addressing the academic and social needs of all students.

> *I think on the positive side, NCLB has really caused people to focus on instruction and to embrace change. On the negative side, it is too lock-step. If we are going to look at data, we need to look at all types of data. And as a qualitative person, data might be anecdotal notes, what do I know about the child? What do I know about the child's family?...That's all data. I cannot just look at a test score, a single measure, and use that to determine a child's fate. And that's where it falls short.*
>
> *We all know what the expectations are, what the bars are that we have to jump over. So, that's the message, that—we know what that is. But we also know we never lose sight of the fact that we teach children. We can't just put a percentage on a child and either highly regard the child or disregard the child. No, they're children, and our primary role is to mold and to care for them, even within the constraints of NCLB.*

In the examples discussed above, the Baker principal and Leadership Team made hard choices about continuing to use teaching, curricular, and assessment strategies that they

believed supported kinds of student learning that would not be supported by systematic adherence to the prescribed Core Curriculum and Benchmark assessments.

An Alternative Metric for Measuring Adequate Yearly Progress: The Value-Added Growth Model

Critics have argued that the status achievement model does not reflect growth in individual student learning over time and that a measure that indicates the value added by the school to individual student achievement would be a more meaningful and equitable assessment.[24] Researchers have also identified several unintended consequences of NCLB's requirements for proficiency by 2014 according to the status achievement metric. These include findings that, in an attempt to make AYP, many strategies for "gaming the system" are used, especially in low-performing schools:

Critics have argued that the status achievement model does not reflect growth in individual student learning over time and that a measure that indicates the value added by the school to individual student achievement would be a more meaningful and equitable assessment.

- Curriculum and pedagogy are increasingly test-driven;
- Test preparation consumes an increasing amount of time;
- Schools focus on students who are on the cusp of becoming proficient;
- Time spent on non-tested subjects has decreased; and
- NCLB sanctions have motivated states to lower standards and develop easier tests.[25]

In response to these concerns, in 2005 the U.S. Secretary of Education authorized a pilot study of "value-added growth models." These models determine AYP by comparing the performance of individual students from year to year, over two or more years. It uses the student as his or her own control and calculates whether the school has made the "expected growth" toward meeting the 2014 proficiency requirements.[26]

---

[24] Among others, Darling-Hammond, 2008; Linn, 2008 argue that a status achievement model does not capture growth in learning over time.

[25] Casserly, M. (2007). America's Great City Schools: Moving in the right direction. In F. Hess, & C. Finn (Eds.), *No remedy left behind: Lessons from the first half-decade of NCLB.* Washington, DC: AEI Press; Darling-Hammond, 2003, 2008; Kornhaber, 2008; Manzo, K. (2008). Analysis finds time stolen from other subjects for reading and math. *Education Week, 27*(25), 6; Mintrop, 2008; Nichols, S. & Berliner D. (2007). *Collateral damage: How high stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press; Poetter, T. (2006). The impact of NCLB on curriculum, teaching and assessment. In T. Poetter, J. Wegwert, & C. Haerr (Eds.), *No Child Left Behind and the illusion of reform: Critical essays by educators.* Lanham, MD: University Press of America; Sunderman, G., Tracey, K., Kim, J. & Orfield, G. (2004). *Listening to teachers: Classroom realities and NCLB.* Cambridge, MA: Harvard Civil Rights Project.

[26] See, for example, Armour, D. (2006, August 16). "Can NCLB Close Achievement Gaps?" *Teachers College Record.* ID Number: 12667. Retrieved on February 8, 2009 from http://www.tcrecord.org; Darling-Hammond, 2008; Linn, 2003, all of whom prefer value-added growth models to status achievement models. Hoff. D.

In 2002, Pennsylvania began to pilot test its own value-added growth model, the Pennsylvania Value Added Assessment (PVAAS). Although Pennsylvania proposed its model for inclusion in a U.S. DOE pilot project, the model was not approved until 2009.[27] In 2007, for the first time, however, Pennsylvania reported the results of its alternative value-added growth model for all school districts in the state. In 2007, 70 district schools were in Corrective Action II. Of these schools, 11 would have made AYP if PVAAS had been approved. Baker was one of these schools, a significant achievement for the school, because it validated the school's commitment to support the learning of *all* students.[28]

Research has shown that using a value-added growth model generally does not enable a school to make AYP if it failed to do so using the status achievement model.[29] In Baker's case, however, it did. Baker's accomplishment according to PVAAS confirmed my observation that rather than labeling it as a "failing school," it could be considered an exemplary one in important respects. According to PVAAS, Baker was a school that had contributed to the "expected growth" of all of its students in reading and math. The gains students made "met or exceeded" Pennsylvania's growth standards.

According to PVAAS (the state's value-added growth metric in 2006-07), Baker was a school that had contributed to the "expected growth" of all of its students.

I would argue that the emphasis that Baker put on the growth of every student, in combination with other factors discussed above in "Establishing a Learning Community for All" interacted to create a robust learning community. While my case study of Baker does not allow me to identify the contribution of specific variables/factors to growth in student achievement on the PSSA, I believe that qualitative case study research, which seeks to tell a coherent, complex story, enables me to offer a credible explanation for the growth in student learning documented by PVAAS.[30]

(2007). Growth models gaining in accountability debate. *Education Week, 27*(16), 22-25 reports that value-added growth models are gaining acceptance as a means for determining AYP.

[27] In 2007, the U.S. Department of Education (DOE) informed Pennsylvania that it would not be permitted to be included in the pilot study and "incorporate a value-added growth model into its accountability system at the present time." While the U.S. DOE did not take issue with the state's basic model for measuring growth, it had concerns about allowing the state to use both a growth model and the state's current performance index (PPI). It encouraged Pennsylvania to submit a revised proposal. *Email to PAC-TE members from John Johnson. (PAC-TE Tidbits July, 2007)* On January 13, 2009, the U.S. Secretary of Education approved a modified version of Pennsylvania's value-added growth model as a way to achieve AYP. Chute, E. (Jan. 13, 2009). New rules credit underachieving students for making progress. Pittsburgh, PA: *Pittsburgh Post Gazette;* Pennsylvania Department of Education. (Jan. 12, 2009). Pennsylvania model for tracking achievement receives federal approval. www.pde.state.pa.us. The model was based on somewhat different methodological assumptions than the original model the state submitted. Arguments made in this paper about the value and fairness of using a value-added growth model for holding schools accountable still stand. However, the recent approval of PVAAS affects the discourse around Philadelphia schools in Corrective Action II that would now make AYP using PVAAS.

[28] It was significant, because of the 11 schools in Corrective Action that would have achieved AYP by PVAAS, Baker had the most grades (K-8), the most subgroups and almost 700 students.

[29] Klein, A. (2007). Impact is slight for early states using growth model. *Education Week, 27*(16), 24-25; Samuels, C. & Davis, M. (2006). Growth model wouldn't have made much difference in schools that missed AYP by status model. *Education Week, 25*(38), 27-28.

[30] Qualitative researchers argue for the value of case studies to analyze complex, context-based realities in settings such as schools. Bassey, M. (1999). *Case study research in educational settings*. Buckingham: Open University;

Policy Implications and Recommendations

Baker's success according to the PVAAS metric, but not according to the status achievement model, raises a number of interrelated policy issues for low achieving schools with robust capacity that show growth, but growth that is not sufficient to meet standard NCLB proficiency targets.

Baker's success according to the PVAAS metric, but not according to the status achievement model, raises a number of interrelated policy issues for low achieving schools with robust capacity that show growth, but growth that is not sufficient to meet AYP the traditional way.

Despite Baker's achievement according to PVAAS in 2007, it was not clear that the school would be able to continue to succeed according to this metric, because the Department of Education still requires schools in states using value-added growth models to be on track to achieve 100 percent proficiency by 2014. The 2014 deadline presented a challenge for Baker because the climb to 100 percent proficiency was steep. Thus, using a value-added growth model could leave Baker in the accountability fix it was in before PVAAS was approved for use in all Pennsylvania school districts.[31]

What changes in the reauthorization of NCLB could create an accountability model which would provide incentives for low-performing schools like Baker to continue improving the learning of every student without imposing undue sanctions?

The three proposals outlined below would address many of the issues discussed in this paper.

1. *All* schools—not just low-performing schools—should be allowed to use multiple measures, including a variety of performance assessments as well as standardized tests, for accountability purposes.[32] Baker's success on PVAAS is based on math and reading scores on the PSSA, a standardized test. Despite this success, members of the school's leadership team and many faculty members did not believe standardized tests alone adequately captured the extent of individual student knowledge and abilities. They believed that using multiple measures, including authentic performance assessments, would be a better and fairer way to assess student learning and thus to hold the school accountable.

2. Among schools that have not achieved AYP, differentiate the sanctions for schools that are showing improvement from those for schools not showing improvement. Such differentiation would mitigate the stress felt by schools that did not meet AYP proficiency levels but had shown growth in student achievement. This proposal also could reduce the stigma that such schools feel

---

Merriam, S. (1988). *Case study research: A qualitative approach*. San Francisco, CA: Jossey–Bass; Yin, R. K. (1994). *Case study research: Design and methods. (2nd Ed.)* Newbury Park, NY: Sage.

[31] See discussion by Sokola, D.,Weinberg, H., Andrzejewski, R. & Doorey, N. (2008). Fixing the flaw in the growth model. *Education Week, 27*(38), 26-27, 29; Weiss, M. (2008). Growth model pilot is not what you think it is. *Education Week, 27*(42), 28-29.

[32] Baker, 2007; Darling-Hammond, 2008; Hawley, 2008; Linn, 2003; and Nichols and Berliner, 2007 all discuss the need for multiple measures to meet AYP.

because the value the school added to student performance would be validated. This more refined categorization would provide useful information to states, districts and the public about ways in which these schools are succeeding and failing. Finally, states would be able to more appropriately tailor sanctions and supports to meet the needs of these schools, if they did not have to rely solely on data about a school's proficiency levels. Variations of this proposal include:

    a. Allow states to replace the status model with a growth model and design a two-tier system of supports and sanctions for schools that do not achieve AYP. Designate schools that have not achieved AYP but have shown growth for "focused improvement" and schools who have not achieved AYP and have not shown growth for "comprehensive improvement."[33]

    b. Use the results of both the status achievement and value-added growth models and apply more severe sanctions to schools that do not achieve AYP according to either model and lesser sanctions for schools who achieve AYP in one.[34]

3. Allow the use of value-added models for holding schools accountable rather than value-added growth models approved by the federal government.[35] Value-added growth models resemble value-added models in some respects. They indicate individual growth over time. Value-added approaches, used by many educational researchers to show growth in student achievement, however, do not require arbitrary time limits or require all schools to make a uniform amount of growth. The expectation that all students make at least a year's growth is reasonable. [36] Moreover, for the purpose of holding schools accountable, a benefit of value-added models similar to value-added growth models, is that they take into consideration out-of-school factors such as social class, race/ethnicity, students with special needs and/or levels of funding that can affect a student's initial achievement. This proposal would alter the meaning of accountability to some extent. However, it would still remain a measure of accountability.

Each of these three proposals, or combinations of them, would benefit low-performing schools, like Baker, whose goals are to improve the performance of *all* students. Because of initial low scores, many of these schools cannot meet yearly AYP targets based on the status achievement model and the requirement that proficiency must be measured by a single score on a standardized test. The proposals discussed above would emphasize

---

[33] Education Trust. (2007). *Recommendations for NCLB Reauthorization.* Washington, DC: Author.

[34] Toch, T. & Harris, D. (2008). Salvaging accountability. *Education Week, 28*(6), 30-31.

[35] Among others, Darling-Hamond, 2008; Koretz. D. (2008). The pending reauthorization of NCLB. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press; and Weiss, 2008, propose using some type of value-added assessment for holding schools accountable.

[36] See, for instance, Koretz, 2008; Booher-Jennings, J. (2005). Below the bubble: Educational triage and the Texas Accountability System. *American Educational Research Journal, 42*(2), 231-268; Toch, T. (2006). *Margin of error: The testing industry in the No Child Left Behind era.* Washington, DC: The Education Sector.

accountability for constructive educational efforts and allow the leadership in schools like Baker to make educational choices they believe are in the best interests of all students.

Conclusion

Baker's story raises significant issues for educators and policy makers concerned with accountability under NCLB. How do we fairly assess schools that appear to be robust learning communities and that improve the performance of all students, but do not meet proficiency targets on a standardized state test?

How do we fairly assess schools that appear to be robust learning communities and that improve the performance of all students, but do not meet proficiency targets on a standardized state test?

Baker provides an example of a school where the press to make AYP according to the status achievement model resulted in some complicated choices for the school. In an era of high-stakes accountability, Baker's leaders struggled to maintain certain educational beliefs and practices. Although they strove to make AYP, they did not let achieving this goal become an end in itself.

The fact that an insufficient percentage of Baker students had achieved AYP proficiency levels according to the status achievement model in 2007, masked the performance of a school that contributed importantly to improved learning of *all* students.

Baker's story testifies to the school's strength as a learning community that supported the learning of each student in the years I did field research at the school. This genuine achievement should be acknowledged by state and district officials. But most of all, it should be celebrated by the school's staff and parents, who could be assured that Baker was not a "failing school" in 2007.

Post Script:

In the 2007-08 and 2008-09 academic years, Baker did not make AYP using either the state's status achievement or value-added growth models. The fact that in fall, 2007 the principal of Baker moved to another district school, along with four key members of the school's Leadership Team, may have contributed to these results.[37]

---

[37] Christman, et. al., 2009, discuss effective school leadership as a key factor in district schools that improve achievement on the PSSA.

References

Armour, D. (2006, August 16). "Can NCLB Close Achievement Gaps?" *Teachers College Record.* ID Number: 12667. Retrieved on February 8, 2009 from http://www.tcrecord.org.

Annenberg Institute for School Reform. (2004*). Instructional coaching: Professional development strategies that work.* Providence, RI: Brown University.

Baker, E. (2007). *The End(s) of Testing.* Presidential Address. Chicago, IL: American Educational Research Association Meetings.

Balfanz, R. & Legters, N. (2008) NCLB and reforming the nation's lowest performing schools. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Bassey, M. (1999). *Case study research in educational settings.* Buckingham: Open University.

Black, P. & Wiliam. D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy and Practice, 5*(1), 7-74.

Booher-Jennings, J. (2005). Below the bubble: Educational triage and the Texas Accountability System. *American Educational Research Journal, 42*(2), 231-268.

Bravmann, S. (2004).Two, four, six, eight, let's all differentiate! Differential education: Yesterday, today and tomorrow. *New Horizons for Learning.*

Casserly, M. (2007). America's Great City Schools: Moving in the right direction. In F. Hess, & C. Finn (Eds.), *No remedy left behind: Lessons from the first half-decade of NCLB.* Washington, DC: AEI Press.

Cech, S. (2008). Test industry split over 'formative' assessment. *Education Week, 28*(4), 1,15.

Chapius, S. & Stiggins, R. (2002). Classroom assessment for learning. *Educational Leadership*, *60*(1), 40-43.

Christman, J. (Spring, 2002). *Baker middle school.* Philadelphia, PA: Consortium for Policy Research in Education & Research for Action.

Christman, J., Neild, R., Bulkley, K., Blanc, S., Liu, R., Mitchell, C. & Travers, E. (2009). *Making the most of interim assessment data: Lessons from Philadelphia.* Philadelphia, PA: Research for Action.

Chute, E. (Jan. 13, 2009). New rules credit underachieving students for making progress. Pittsburgh, PA: *Pittsburgh Post Gazette.*

Coalition of Essential Schools Resources, www.essentialschools.org.

Darling-Hammond, L. (2003, February 16). *Teachers College Record.* ID Number: 11109. Retrieved on February 8, 2009 from http://www.tcrecord.org.

Darling-Hammond, L. (2008). Improving high schools and the role of NCLB. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Darling-Hammond, L., Ancess, J. & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press

Darling-Hammond, L. & McLaughlin, M. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(8), 597–604.

Education Trust. (2007). *Recommendations for NCLB Reauthorization.* Washington, DC: Author.

Elmore, R. (2000). *Building a new structure for school leadership.* Washington, DC: The Albert Shanker Institute.

Fullan, M. & Hargreaves, A. (1991). *What's worth fighting for in your school?* New York: Teachers College Press.

Fuller, B., Gesicki, K., Kang, E. & Wright, J. (2006). *Is the No Child Left Behind Act working? The reliability of how states track achievement.* Berkeley, CA: Policy Analysis for California Education.

Hawley, W. (2008). NCLB and continuous improvement. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Herlihy, C. & Kemple, J. (2004). *The Talent Development middle school model: context, components, and initial impacts on students' performance and attendance.* New York: MDRC.

Hess, F. & Finn, C. (2007). Conclusion: Can this law be fixed? A hard look at the NCLB remedies. In F. Hess and C. Finn (Eds.), *No remedy left behind: Lessons from the first half-decade of NCLB.* Washington, DC: AEI Press.

Hoff. D. (2007). Growth models gaining in accountability debate. *Education Week, 27*(16), 22-25.

Inger, M. (1993). Teacher collaboration in secondary schools. *Center Focus, 2.*

Jacobs, E. (January 13, 2009). NCLB: Recipe for success or failure? Washington, DC: *D.C. Examiner.*
.
Klein, A. (2007). Impact is slight for early states using growth model. *Education Week, 27*(16), 24-25.

Koretz. D. (2008). The pending reauthorization of NCLB. In G. Sunderman (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Kornhaber, M. (2008). Beyond standardization in school accountability. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Lambert, L. (2002). A framework for shared leadership. *Educational Leadership, 59*(8), 37-40.

Lather, P. (1991*). Getting smart: Feminist research and pedagogy with/in the postmodern.* London: Routledge, Chapman and Hall, Inc.

Laurence-Brown, D. (2004). Differentiated instruction: Inclusive strategies for standards-based learning that benefit the whole class. *American Secondary Education, 32*(3), 34-62.

Lee, J. (2008). Two takes on the impact of NCLB on academic achievement. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Lewis, A. (2002). A horse called NCLB. *Phi Delta Kappan,* 84(13), 179-180.

Lieberman, A. (1990). *Schools as collaborative cultures: Creating the future now.* Bristol, PA: The Falmer Press.

Linn, R. (2003). Accountability, responsibility and reasonable expectations. *Educational Researcher, 32*(7), 3-13,

Linn, R. (2008). Toward a more effective definition of AYP. In G. Sunderman, (Ed.), *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press**.**

Little, J.W. (1987). Teachers as colleagues. In V. Richardson-Koehler (Ed.), *Educators' handbook: A research perspective.* New York: Longman.

Mac Iver, D. & Balfanz, R. (2000). Transforming high-poverty urban middle schools into strong learning institutions: Lessons from the first five years of the Talent Development middle school. *Journal of Education for Students Placed at Risk, 5,* 1- 2.

Manzo, K. (2008). Analysis finds time stolen from other subjects for reading and math. *Education Week, 27*(25), 6.

Marion, S., White, C., Carlson, D., Erpenbach, W., Rabinowitz, A. & Sheinker, J. (2002). Making valid and reliable decisions in determining AYP. *Implementing the state accountability requirements under the No Child Left Behind Act of 2001.* Washington, DC: Council of Chief State School Officers.

McGuinn, P. (2007). New Jersey: Equity meets accountability. In F. Hess and C. Finn, (Eds.), *No remedy left behind: Lessons from the first half-decade of NCLB.* Washington, DC: AEI Press.

Merriam, S. (1988). *Case study research: A qualitative approach.* San Francisco, CA: Jossey-Bass.

Mintrop, H. (2008). Low-performing schools' programs and state capacity requirements. In G. Sunderman, (Ed.), *Holding NCLB Accountable: Achieving Equity, Accountability and School Reform.* Thousand Oaks, CA: Corwin Press

Neill, M. (2003). Leaving children behind: How No Child Left Behind will fail our children. *Phi Delta Kappan, 85*(3), 225-228.

Neufeld, B. & Roper, D. (2003). *Coaching: A strategy for developing instructional capacity, promises, and practicalities.* Providence, RI: Annenberg Institute for School Reform.

Nichols, S. & Berliner D. (2007). *Collateral damage: How high stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press.

Novak, J. & Fuller, B. (2003). Penalizing diverse schools? Similar schools, but different students bring federal sanctions. *Policy Brief.* Berkeley, CA: University of California.

Owens, A. & Sunderman, G. (2006). *School accountability under NCLB.* Cambridge, MA: Harvard Civil Rights Project.

Pennsylvania Department of Education. (Jan. 12, 2009). Pennsylvania model for tracking achievement receives federal approval. www.pde.state.pa.us.

Poglinco, S., Bach, A., Hovde, K., Rosenblum, S., Saunders, M. & Supovitz, J. (2003). *The heart of the matter: The coaching model in America's Choice schools.* Philadelphia, PA: Consortium for Policy Research in Education.

Poetter, T. (2006). The impact of NCLB on curriculum, teaching and assessment. In T. Poetter, J. Wegwert, & C. Haerr (Eds.), *No Child Left Behind and the illusion of reform: Critical essays by educators.* Lanham, MD: University Press of America.

Rothstein, R., Jacobsen, R. & Wilder, T. (2006), Proficiency for all: An oxymoron. *Education Week, 26*(13), 32, 44.

Rothstein, R. & Jacobsen, R. (2006). The goals of education. *Phi Delta Kappan, 88*(4), 264-272.

Samuels, C. & Davis, M. (2006). Growth model wouldn't have made much difference in schools that missed AYP by status model. *Education Week, 25*(38), 27-28.

Sokola, D.,Weinberg, H., Andrzejewski, R. & Doorey, N. (2008). Fixing the flaw in the growth model. *Education Week, 27*(38), 26-27, 29.

Spillane J., Halverson, D. & Diamond, J. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23-28.

Stacey, J. (1998). *Can there be a feminist ethnography?* Amsterdam: Elsevier Science, Ltd.

Stiggins, R. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement, Issues and Practice, 20*(3), 5-15.

Sunderman, G., Tracey, K., Kim, J. & Orfield, G. (2004). *Listening to teachers: Classroom realities and NCLB.* Cambridge, MA: Harvard Civil Rights Project.

Sunderman, G. (2008). Introduction. *Holding NCLB accountable: Achieving equity, accountability and school reform.* Thousand Oaks, CA: Corwin Press.

Toch, T. (2006). *Margin of error: The testing industry in the No Child Left Behind era.* Washington, DC: The Education Sector.

Toch, T. & Harris, D. (2008). Salvaging accountability. *Education Week, 28*(6), 30-31.

Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners.* Association for Supervision and Curriculum Development.

Tomlinson, C. (2004). *How to differentiate instruction in mixed-ability classrooms.* Upper Saddle River, NJ: Prentice Hall.

Tracey, C., Sunderman, G. & Orfield, G. (2005*). Changing district accountability standards: Implications for rural equity.* Cambridge MA: Harvard Civil Rights Project.

Travers, E. & Blanc, S. (Jan. 8, 2007). School yardstick fails to measure up. Philadelphia, PA: *Philadelphia Inquirer,* B, 2.

Weiss, M. (2008). Growth model pilot is not what you think it is. *Education Week, 27*(42), 28-29.

Western Regional Educational Laboratory. (2000). *Teachers who learn, kids who achieve: A look at schools with model professional development.* San Francisco, CA: West Ed.

Wiggins, C. (1998). *Educational assessment: Designing assessments to inform and interpret student performance.* San Francisco, CA: Jossey-Bass.

Wiliam, D. (2001). An overview of the relationship between assessment and the curriculum. In D. Scott (Ed.), *Curriculum and assessment.* Greenwich, CT: JAI Press.

Wiliam. D. (June, 2004) *Keeping on track: Integrating assessment with instruction.* Invited Address to the 30th Annual Conference of the International Association for Educational Assessment, Philadelphia, PA.

Yin, R. K. (1994). *Case study research: Design and methods. (2nd Ed.)* Newbury Park, NY: Sage.

RESEARCH FOR *ACTION*

3701 Chestnut Street
Philadelphia, PA 19104
TEL (215) 823-2500
FAX (215) 823-2510
www.researchforaction.org