Students, Teachers and High Standards Reforms:
Negotiating Education Policy, Classroom Practice and Student Outcomes
In Philadelphia Middle Schools

Analytic Essay from a School Reform Planning Grant

Submitted to the Spencer Foundation

By

Jolley Bruce Christman

Thea Abu El-Haj

Rosalie Rolon Dow

Research for Action

and

Ellen Foley

Annenberg Institute for School Reform

Students, Teachers and High Standards Reforms: Negotiating Education Policy, Classroom Practice and Student Outcomes in Philadelphia Middle Schools

I. INTRODUCTION

> We just told them that they needed to do it to graduate. (Principal, Parks Middle School)
>
> [The teachers] don't ever really talk about this stuff [the new promotion requirements]. It's like a big secret and they just give us all the work. They don't tell us why we're doing it. They just tell us it passes, it passes us from eighth grade to ninth grade. (Danielle, eighth grade student at Parks Middle School)

Despite a spate of educational reform efforts over the past three decades, too many students in America's urban schools continue to succeed and fail along well-worn trails of race, class, ethnicity and gender. (Education Trust, 2000; Hornbeck & Ingram, 1998; Jencks & Phillips, 1998; Jerald, 1998; Johnston & Viadero, 2000; U.S. Department of Education, 2001) Standards-based reforms, often accompanied by high stakes promotion and graduation requirements for students and accountability measures for schools, are the current remedy for raising achievement levels of students of color and those living in poverty. It seems common sense that these reforms and their increased requirements demand new actions on the part of students. Yet, as Danielle points to in her comment above, adults have paid scant attention to how young people make sense of what they are now being asked to do and how they respond to these new demands (Corbett & Wilson, 1995; Fullan, 1991; Ozga, 2000).

In this essay we argue for research on school reform that positions students at the center of policy and we propose directions for future investigation. To date, research on high standards reforms has focused on adults, not students, and has offered insights about the myriad instructional changes teachers must make and the complex decisions with which they must contend to translate policy into on the ground practice.[i] But it has not offered policy makers analyses of how students perceive and engage with these policies.[ii]

We also propose research that invites students to examine and deliberate upon educational reform within the context of an interpretive community of researchers, practitioners, parents and policy-makers. The concept of interpretive community draws from the tradition of research that seeks to understand the world by interpreting how social actors construct meaning through talk and action (Schwandt, 1994; Wasser & Bresler, 1996).[iii] Within an interpretive community multiple perspectives exist in tension with one another

offering diverse entry points for understanding the questions: What is happening and what does it mean to different groups? Finally, the idea of "community" sets certain norms for the space in which the interpretation will occur, including trust that different perspectives will be heard and tolerated.

During the research planning process funded by the Spencer Foundation grant, we conducted exploratory research into eighth grade students' experiences as they negotiated their way through the first phase of the Philadelphia's school district's new promotion policy and its accompanying assessments.[iv] We tracked academic outcome data for all eighth graders in the district, and we took a closer look at the experiences of a small group of students in two middle schools. (See Appendix A for a discussion of research methods and background on the two school sites.) We learned that students experience reforms as a "big secret" and, like their teachers, feel buffeted by a myriad of mandates that threaten them with failure. We also learned that they can be astute critics of the contradictions inherent in policies that punish them for the failures of schools and society to provide equitable education for children in urban schools.

During this planning period, we also piloted an approach for involving multiple stakeholders in our research process. Drawing on Research for Action's strong history of collaboration with teachers, parents, students, administrators and community members, we brought together individuals from a broad range of experiences, perspectives and positions at two points during the planning. Our goal was to begin to build a safe community where people could surface emerging issues and questions and critically examine and interpret data that our research team had collected. Participants were hungry for these discussions. Together we gained new insights about students and their experiences with the promotion policy. In addition, practitioners offered us valuable suggestions for shaping the direction of future collaborative research. In the conclusion of this paper we discuss some of the important lessons from these meetings.

In the following section we analyze how students are positioned in the rhetoric on standards-based reform, particularly the assumptions about students in Philadelphia's systemic reform effort. We then report on themes from our exploratory research that speak to these assumptions. Finally, we synthesize lessons from our planning phase to chart a direction for future research.

II – Philadephia's standards-based reform: Assumptions about students

Philadelphia is the seventh largest school district in the United States, and the city has been engaged in standards-based reform for six years. Content standards in seven subject areas outline the knowledge and skills Philadelphia students should acquire, with benchmarks defined at the fourth, eighth and eleventh grades (School District of Philadelphia, 1996). An accountability system rewards schools' progress or sanctions their decline every two years on the Professional Responsibility Index (PRI), a combined measure of five indicators including standardized test scores, teacher and student attendance, and promotion or persistence rates. Further, a promotion policy, adopted beginning in 1999-2000, eliminates social promotion, increases course-taking and passing requirements, mandates multidisciplinary and service learning projects for students at benchmark grades, and requires students to pass the Stanford Achievement Test, 9th Edition (SAT-9).

Philadelphia's set of standards-based reform policies, like similar policies being implemented across the nation, reflects competing discourses about the problems that lead to low student achievement and conflicting means for improving it. The rewards and sanctions in the promotion policy and accountability system mimic corporate performance models that rely on extrinsic, carrot-and-stick incentives. They assume that if the negative consequences of poor performance are increased, students and their teachers will respond with improved effort and achievement.

A mantra of Philadelphia district leadership was: "We must behave as if we believe that all students can achieve at high levels." "It's not good enough to try hard. We must make a difference in what children achieve (School District of Philadelphia, 1995)." These messages embody a second discourse about standards, one that emphasizes the development of high expectations as a policy mechanism for leveraging socially equitable student outcomes. The assumption is that academic performance differences by race and class are the result of adults' beliefs that poor students and students of color cannot achieve at high levels. Changing adult expectations is the critical component of improved student performance. If teachers believe

that students can achieve at high levels, then they will teach an enriched and rigorous curriculum to all students, regardless of their race, class and gender. It is assumed that students' beliefs about themselves and their attitudes toward learning will follow the adult lead.

The district's mandate for project-based curricula aligns with a third discourse underlying standards-based reform in this country: high standards can leverage constructivist pedagogy (see for example, Darling-Hammond, 1997; Wheelock, 1998). Here the assumption is that students are open and ready for new roles such as pursuing their own questions and reflecting on their work.

III – Exploratory Research Findings

Overview of the exploratory study

Our exploratory research focused on eighth grade where changes in promotion requirements were particularly dramatic. For the first time, eighth graders were required to pass all of their major subjects (Reading/English/ Language Arts, Mathematics, Science and Social Studies) and complete a multi-disciplinary project to be promoted to ninth grade.

We decided, for the purposes of the exploratory study, to look at two very different types of assessment incorporated into the promotion policy: standardized tests and proficiency on the multidisciplinary project. This project is intended to help eighth graders demonstrate their competency in more than one subject, exhibit strong writing and research skills, and address at least one of the District's "cross-cutting competencies" — citizenship, communication, multicultural competence, problem solving, school-to-career, or technology. It was the first policy mandate that directly encouraged teachers to adopt constructivist teaching practices.

As originally conceived, the promotion policy also required that by 2001, eighth graders obtain a minimum score of Below Basic III in reading and math on the Stanford Achievement Test 9th Edition (SAT-9) in order to be promoted to ninth grade. Although that policy was not implemented at the time of our fieldwork, the SAT-9 tests already constituted a high stakes test for students. Admission to magnet schools and programs were in large part determined by the requirement to score in the 85th percentile or better in

math and reading.[v] Although it incorporates both open-ended and multiple-choice response formats and is criterion-referenced, the SAT-9 is a much more traditional way of measuring student knowledge and school progress than the project.

In order to get a real, on the ground, sense of how eighth grade students experienced these two different types of assessment, we spent last spring as participant observers in two Philadelphia middle schools. The two schools in the study were not selected to be representative of Philadelphia middle schools. On the contrary, we chose these schools because we were aware that each had a history of engaging with educational reform and had some experience with project based learning. Our intent was not to understand the extent of participation in the reform across the district, but rather to observe teachers and students we knew would be engaged in multidisciplinary projects.

Despite their shared history undertaking educational reform, the two schools were, in fact, very different. (See Tables 1 & 2 in Appendix A for detailed data on the two schools.) Parks is a small school serving approximately 240 students who arrive each day from neighborhoods all over the city. Part of the city's desegregation program, Parks students are admitted through a lottery system that allots places according to a formula for achieving racial diversity. As a small school and a school that serves students from a range of economic and racial/ethnic backgrounds, Parks is not typical for the district. In contrast, Cobb is a more racially isolated school, serving mostly African-American students with a small population of Asian-American and Latino students. Approximately 900 students in grades 6-8 attended Cobb. In 1998-99 90% of Cobb's student population lived in poverty, approximately 8% were linguistically diverse and participated in the school's ESL program, and 15% were special education students.

The two schools' experiences with the SAT-9s were also very different. Parks's practitioners were particularly incensed by the district's reliance on the SAT-9 as the primary measure of school and (under the soon to be implemented promotion policy) individual student performance. From the perspectives of the practitioners we spoke with the test worked at cross-purposes to the kinds of deep learning they hoped the school was promoting. Although the open-ended questions on the exam were more reflective of the types of

learning the school was implementing, the multiple-choice sections of the exam offered, in the words of the principal, “absolutely no correlation with the curriculum goals.” The principal argued, “the SAT-9 is not robust enough to tell you what to teach.” Instead of designing curriculum that would teach to the test, Parks teachers tried to focus some of their instruction around specific test-taking strategies that would be useful to the students. In part this was a luxury afforded the school by virtue of the fact that their students generally tested above the city average in reading and math.

Cobb, on the other hand, was under much greater pressure to improve their SAT-9 test scores because it had not met its accountability targets for the first two-year cycle. A new principal arrived in 1997, determined to “turn the school around,” and multiple reform initiatives ensued, including: adoption of a whole school reform design that provided a core curriculum and professional development for teachers, a myriad of community partnerships, a technology initiative, project-based learning. Not surprisingly, the school also directed considerable attention to preparing students for the SAT-9. The principal made staff assignments with the test in mind, putting the strongest teachers in tested grades and subjects; seventh and eighth graders received several periods of test prep each week in the months preceding the test; and classroom teachers were encouraged to incorporate assignments and tasks that were similar to items on the SAT-9 in their daily instruction.

In what follows, we look inside these two schools to learn about how students made sense of the new promotion policy, particularly the requirement for the multidisciplinary project and the SAT-9.

Inside the two schools

We began our exploratory research with the questions: How do Philadelphia eighth graders understand what is being asked of them in the district’s new promotion policy and two of its accompanying assessments, the SAT-9 and the multidisciplinary project? What factors influence their understanding? What concerns do they have about the promotion policy? How does their understanding match the intentions of the policy? What we observed and heard from students and their teachers suggests that policymakers' attempts to mandate higher standards from the top down seriously underestimates the classroom “reculturing” (Fullan, 1991) necessary for students to become invested enough in their school work to

engage in the revision, refinement and reflection that lead to deep and substantive learning. Yet, we also saw moments of possibility in which students discovered meaning in work connected to real-world audiences and purposes and, with guidance from their teachers, put forth good effort to do their best work. Students also spoke passionately and critically about the promotion policy, pointing to its political underpinnings and its incapacity to address fundamental inequities in the system.

Below we focus our analysis primarily on how Philadelphia's new multidisciplinary project requirement was implemented in the two schools and students' responses to that implementation. We turn next to students' critiques of the promotion policy, especially their analysis of the testing component. In our exploratory research we collected a range of other data on students' experiences with the promotion policy.[vi] Our focus for this essay reflects our desire to open up compelling questions we now have about how classroom interactions between students and teachers do (or do not) invite students to be partners in their learning in new ways and how students' observations about school reform might offer new insights for adults in the business of research and reform.

Multidisciplinary Projects at Cobb Middle School

One of the many initiatives the new principal brought to Cobb was project-based learning. The principal described the thinking behind "Project Day" which was instituted in 1998-99 and occurred once a month:

> Our kids need the chance to do exciting things. We need to get them engaged with learning, give them the chance to be active learners. That's why I'm pushing project-based learning. We started this before the District made it a priority.

Thus, Cobb tackled the District's mandate for eighth grade multidisciplinary projects with some experience under its belt.

We focused our research at Cobb in two eighth grade Reading/English/Language Arts classrooms—Ms. Foster's class and Ms. Wald's class. In Ms. Foster's class, students were developing exhibits for the annual health fair held by their small learning community (SLC).[vii] The teachers in Ms. Foster's SLC, Health, Technology and Wellness, chose to have students' work on the Health Fair count as their eighth grade

project. Eighth graders identified a topic, such as AIDS, peer pressure, the cardio-vascular system, and then worked independently, in pairs, or groups to prepare their exhibits. They developed essential questions such as “Where did AIDS come from?” “What is peer pressure?” “What do teens need to know about hygiene?” to guide their research.

The curriculum in Ms. Wald’s classroom was almost entirely project-based. For example, students were simultaneously producing the art and text for an “Alphabet Book” as part of their study of the Harlem Renaissance and conducting an inventory of neighborhood buildings. In the former project, each student was responsible for one letter of the alphabet. The production of this drawing actually involved several steps. Students imitated the style of several Harlem Renaissance artists and then chose a style for drawing their letter. Then, students chose a word that began with the letter they were responsible for and which was related to the civil rights movement (e.g., B is for “Bus Boycott”). They used that word in the illustration of their letter. When all the illustrations were complete, they were bound and published in a book. An urban landscape architect and a school district employee whose job it was to promote community involvement in schools collaborated on a second project in Ms. Wald’s class. Students toured the neighborhood around the school, noting which buildings were residences, businesses or abandoned. They entered data from this neighborhood inventory in a spreadsheet and mapped them. They analyzed the data and prepared Power Point presentations for the mayor and other community leaders on their findings.

Multidisciplinary Projects at Parks Middle School

Multidisciplinary projects were not entirely new to the students and staff at Parks either. In the late 1980s, Parks joined the Coalition of Essential Schools and became known in the district as a place where practitioners, in collaboration with parents and students are routinely involved in the re-examination of their educational practices. Within this context of reform, project-based learning has been an increasingly important component of the curriculum.

As its turns out, Parks's prior experience with multidisciplinary projects did not translate into easy implementation of the district's new project requirement. Most of the eighth grade team had had little or no experience with project-based learning. During the fall term the team made little progress in developing a

multidisciplinary project for the eighth graders other than choosing a focus, the Civil War, and an essential question: What is conflict? Finally in December, the principal pressured the eighth grade team to select a pre-existing project about the Civil War from a web site. She did so because she feared that students would be denied entrance to ninth grade based on a poor grade on a project that they had not had adequate time to complete.

Students chose from several different books set before or during the Civil War and then responded to a series of short activities that were meant to be generalizable to every book.[viii] Some sample activities included: create a "Wanted poster" for the villain; explain how a character felt at the beginning middle and end of the story; and create a restaurant menu appropriate for that time period. In their Social Studies class, students were also asked to respond in at least a full written page to journal prompts about conflict and to include the journal in their project folder. However, in the end, projects submitted without a journal were rated as proficient. Although the Language Arts teacher took the primary responsibility for this project, all the team's teachers provided class time for students to work on their projects. The schedule was rearranged frequently to give students the opportunity to work on the activities; they completed the entire project in school.

Students' Perspectives: Cross-cutting themes

Although the projects we observed at Cobb and Parks were very different, themes emerged across the sites.

Theme 1: In many instances, the multidisciplinary project encouraged student compliance and attention to form, rather than engagement with process, content, and quality of work.

At both Cobb and Parks, we saw evidence that the focus of the projects was on form, not content. Students and teachers did not emphasize learning goals of deeper knowledge or increased understanding of the topic, but rather focused on organization, appearance and compliance. A Cobb student told us:

> We going to be graded on creativity, appearance, the way you present yourself, the way you present your project, the ways it's set up. Stuff like that. So, basically, you just going to be graded on what you should already know.

We noted a similar emphasis in the projects done by the Parks eighth graders. Over and over, students talked to us about numbers—how many pieces of work needed to be in their project portfolio, how many pages they were required to write—and about format—the project had to have a title page and clearly note what activities the student had completed. As to the larger purpose students imagined the projects served, many described them as exercises in getting them organized. As Crystal, a student at Parks, put it:

> They're just trying to teach us about how we organize our work. That's exactly what we're doing. We have a folder. We just do certain – they give us these little colorful sheets [referring to the sheets on which the activities were listed] and two blue, two yellow, two green and there was one pink sheet for our cover page and you could do the first. The first blue one was knowledge and then you pick a number from there, an activity number and then you would just go from there.

Further, an overemphasis on form sometimes led to missed opportunities for students to develop skills. One example comes from a researcher's field notes from Cobb:

> Two students' work on hygiene revolved around soliciting companies for samples of their products. When they got ready to write the letter, their teacher told them, "I have a letter that you can use." Danielle replied, "I want to write my own letter." "No, just use this one. It will save time."

Theme 2: Students complained about work they considered "easy" but were ambivalent about taking on more challenging assignments.

Students resented work they viewed as "easy" or in middle school parlance, "stupid." Students, like Marina, a Parks student, resented the low expectations:

> But it's like, I thought they were making a big deal like, "If you don't do this project, you're not graduating." I thought it was going to be hard. I thought it was going to be like something that we really -- My little cousin, that's in like, in first grade can do it. I'm not complaining, I'm not complaining. It's just that they were making a big deal about something so little.

At both schools, students told us they desired work that made them "think" but they were also ambivalent about investing more in their schoolwork. However, as reflected above in Marina's statement, students were not quite sure they wanted to "complain" about how easy the work was. At times students did complain about requirements that seemed to demand more substantive work, for example the requirement to "write a full page" in response to journal prompts. In some cases these complaints revealed that the requirement was a real challenge; in others, they articulated students' observations that teachers were ratcheting up

expectations for quantity of work, but not for clearer or deeper thinking. For example, Frank told us he felt forced to write "filler" and Dan complained that,

> We had to do like a page, like you had to just write, it was like… you should just be able to write it. Just, just like writing as much as you, like, want… 'cuz it's then you trying to get to a page and that means you writing stuff over [again] if you're going to [get to a full page.]

Theme 3: Students were more invested in the projects when they knew their work had a real-world purpose and audience; in other cases they merely completed basic project requirements with minimal intellectual and emotional engagement.

At Cobb, almost all of the students we interviewed were eager to show us their work and we hypothesize that their engagement stemmed from the real-world usefulness of the work they were asked to do. In the health fair project, the fact that a panel of teachers would assess their exhibits communicated a sense of importance to the students. Jamal seemed excited that he would be able to share what he had learned with others.

> So, I wrote like two pages about STD and AIDS but I got a lot more to write. And when I get done my writing, I'm gonna make a brochure about AIDS, um and I'm a pass'em out at the health fair.

The students also felt empowered by learning experiences that offered them a chance to have their voices heard in the community. Diona told us:

> I learned that mostly it's the people that are in charge that have more say in the government because they're the ones that step out and do something about how they feel and sometimes when somebody doesn't agree with their opinion, well instead of just, you know, just talking, they should at least do something to change, to change the other person's opinion [thumping the desk for emphasis]. So that they can make an opinion too.

Mary, another Cobb student, agreed with Diona:

> And if we, if we do show up at meetings and things, and tell them and tell the mayor what we need to do in our communities, we can get it fixed. It's not going to be done in a month or the same day that you say it's going to be done, but it will be done.

At Parks, the eighth grade team had chosen a topic rich with possibilities for connection with students' lives, yet few students discovered any relevance. The project had been presented as a study of conflict, with the Civil War as the example. However, students noted that attempts to draw connections

between the Civil War and the broader idea of conflict were tenuous at best. David thought he would have an opportunity to think about conflict resolution, but instead he saw the project as pointless.

> I don't see what's the point of it. They're teaching us on conflict, about like, they want us to stop like fighting or something like that about the Civil War. Civil war is just probably a topic. I don't think it's really a point to that. It's not like stopping us from arguing, fighting and this other stuff that we do.

In a Parks classroom not described in this paper, we observed similar lack of investment, even when project topics were student-selected and focused on content that we would have predicted to be compelling to students. The distance that many students felt from topics that adults had imagined would build on students' life experiences and interests suggests that practitioners and researchers need deeper understandings of what would constitute effective culturally relevant curriculum (Ladson-Billings, 1994).

Theme 4: Students needed adult guidance. Their efforts at the multidisciplinary projects were fragile and highly sensitive to their teachers' uncertainty about: what projects entailed; how to structure projects so that they were doable, but not oversimplified; how to coach students to deep conceptual understanding.

Students in both schools needed and appreciated consistent guidance and clear examples about what their projects should entail. To help keep students "on the point," Keyanna, a student working on the Health Fair project at Cobb, told us:

> We also have 'Reflections' which tells what you did so far. So like I have Step 1 and Step 2 and Step 3 and Step 4 [to complete the whole project] and so far I did step 1 and 2. It's helpful because when I was working with other people in my group all the different things people was doing kept getting confused.

But such structures and coaching were not always applied in ways that helped students develop deeper conceptual understanding of subject matter. For example, some of the essential questions students developed for their health fair projects did not lend themselves to rich intellectual exploration.

At Parks, students were baffled when there were drastic changes in project parameters resulting from the behind the scenes negotiation between the principal and the eighth grade teachers. Alyse expressed

her irritation with her teachers, “They confused. I got really confused about what it was.” Even without such adult missteps, students expressed frustration. In Ms. Wald’s class, where curriculum was frequently organic and constructivist—emerging from students’ (and the teacher’s) interests, students were puzzled by pedagogy that diverged from their previous experiences in school. In this exchange Yung kept completing Robin’s sentences as they described their classroom:

> Robin: Ms. Wald she is, she’s unorganized ‘cuz one moment she gives us a project to do and then in the next moment like…
>
> Yung: She gives us another one.
>
> Robin: A month or a week, next week, she gives us a different project and then she forgets about the project she assigned before.
>
> Yung: She cancelled it.
>
> Robin: And then she just doesn’t get, like, what she’s supposed to do straight, so, it’s kind of fun because she does give a lot of reports… it’s fun when you do a lot of reports, especially with a group or something.

Students had fun in Ms. Wald’s class despite finding her teaching style and more organic approach to curriculum “disorganized.” Indeed, researchers observing her classroom found the scope and underlying vision of the curriculum both impressive and somewhat scattered. Still, Ms. Wald succeeded where other teachers faltered. In another classroom we observed, but that is not described in this paper, teachers asked students to take on projects with a similar scope to those Ms. Wald’s class did, but struggled to give students the coaching they needed to engage in and successfully complete the work. A researcher’s field notes from the teachers’ team meeting described the challenges they faced:

> The teachers were in agreement that the students didn’t know what to do and that the students needed more guidance. The math teacher, in particular, was advocating breaking down the assignment for the students. She felt that the teachers needed to tell the students to include an introduction, for example, then tell them what that would entail, what questions they would have to answer in the body of the report and what a conclusion would entail, etc. She [said] that students were complaining that they didn’t understand what they were supposed to do. The other teachers seemed to agree but not as vehemently. [Another] teacher proposed giving students more class time to do the work.

Ms. Wald, on the other hand, seemed to be able to provide the support necessary for students to take on and complete challenging, comprehensive projects without resorting to a recipe. In part, her success seemed to stem from her emphasis on continuous revision. As Shakira told us:

> [The teacher] will not let you go away if the paper is not perfect. She gave me an A on my paper, but she said she wanted me to write more, make it better. I was like I got an A already. Ain't nothing past an A.

Later, when she was asked whether the continual revisions required by her teacher helped her learn about writing, Shakira responded:

> I learned about it, and I'm learning. Well, it's a gift… Sometimes I think of it as a gift some times because we learned the stuff that in our other classes… [My teacher is] the only teacher who teaches that. Other teachers teach English and grammar and stuff like that, but [my teacher is] worrying about paragraphs and sentences and indentation and stuff like that. And it's too much and we're the only class that's getting this treatment.

Again, Shakira reveals her ambivalence about Ms. Wald's teaching style: "it's a gift" and "it's too much."

Students' perspectives on their experiences with the multidisciplinary projects raised many questions that we believe bear further investigation including: *Under what circumstances do students reconstruct their roles in classrooms to be more active and invested learners who routinely revise and refine their work to higher standards? What does this reconstruction look like? What do students say about it? What features of the classroom and school context support it? What constitutes effective academic scaffolding that helps students develop new skills and knowledge?* In urban settings these questions must include attention to cultural relevance. Research on students of color and those living in poverty shows the need for education that allows them to choose academic success while supporting their identities and connections to their communities (Delpit, 1995; Fordham and Ogbu, 1986; Ladson-Billings, 1994). *What are students' perspectives on culturally relevant teaching? What is 'culturally relevant' to students?* (Other work like that of Ladson-Billings addresses similar questions from teachers' perspectives.)

Theme 5: Students knew their schools needed improvement and speculated that such improvement was the goal of the new promotion policy. However, they were angry about a

system that used a standardized test measure to determine access to educational opportunity (admission to magnet programs) and to indicate academic success or failure. They asked why they should be held accountable for the failure of schools to educate them.

When we asked students to tell us why the project mandate and other new promotion requirements had been put in place, they speculated that stiffer promotion requirements were instituted to make schools better. As Emma, a student who had recently transferred to Parks from a suburban school, put it,

> I think maybe the point [to the new requirements] was there's a lot of bad public schools in our city. Actually, most of the public schools are not good schools. And I think this was supposed to be a way to make people learn more before they graduate.

Students viewed the new promotion requirements as a way for schools to improve by requiring students to earn passing grades through more effort and work. Kareem told us:

> Probably [the policy is] because people ain't going to school, so they're making it harder because people passing without doing nothing, and they made it so that people got to do their project, they got to earn their pay, they got to earn their passing grades, I guess. Because a lot of people is getting out of, is graduating out of high school and don't know nothing. They just graduating.

Even as students viewed the intention of new promotion policies as leverage for improving their schools, they were cynical about the ability of the new reform policies to provide fair opportunities for all students. Students' comments on the motives of administrators and politicians for the reforms do not include any discussion about how these powerful individuals may be considering the students' best interests. Isabelle, a Parks student, stated that perhaps the district superintendent "decided he would do one more thing before he died" while Emma saw the policy as an outcome of new city administration and a mayor who "changed a whole bunch of rules around here."

Students' sense of unfairness was most pronounced in their critiques of the use of the SAT-9. Many students were particularly angered by the unfairness of a system that placed increasing weight on standardized test scores as an indicator of an individual's academic success and future potential. The students we interviewed already had a flavor for the high stakes nature of this test. Seventh grade scores were used to determine admission to several of the district's most prestigious high schools. Many students believed that their long-term performance in school, as measured by grades, attendance and behavior should far outweigh

their test scores. Students argued against the use of single test scores stating, "you could be having a bad day" or "some people are not good at taking tests, but some people are good at other stuff." Furthermore, students were frustrated by the lack of connections between the test material and what they were learning in school.

The students we spoke with also raised questions about the fairness and validity of using tests as a measure of their ability, rather than as an indicator of the education they had received. Trying to make sense of where the accountability ax should fall Marina said,

> Our teacher was saying that the test shows them how much the teachers are teaching us, but if that's like part of it, why are they even showing us what high school we're good enough to go into? So it's not our fault that we're doing like, having those test scores.

If the purpose of the test is to uncover the effectiveness or ineffectiveness of teachers' instructional ability, why should students pay a price for low achievement? Ashley, an African-American girl, put it this way:

> The SAT-9 should be to see if you need extra help cause you don't know something. And it should tell people how the whole school is doing to see if the school is doing a good job or not. So parents and everybody will know if their kids are in a good school. But not for whether we pass and all that.

Students were critical of and, in some instances, cynical about the new promotion requirements. They argued with policies that, for example, exclude them from the district's few "good" high schools, reduce their performance history to a single test score, and fail to provide the supports necessary for them to do well. They also raised new questions for our research, such as: *From students' perspectives how do new promotion requirements speak to (or overlook) issues of fairness and equity? In turn, how do students' perceptions of the purposes and fairness of the policies contribute to (or constrain) their willingness to engage with new requirements and become more active and invested learners?*

IV -- Conclusion

We agree with Corbett and Wilson (1995) that putting students at the center must be the next task of educational research and reform. Our exploratory research reinforced our sense of the urgent need to scrutinize how high stakes reform policies are playing out in particular local settings for students, especially poor students and students of color. These reforms continue to be touted as the silver bullet for leveraging

improved student outcomes, even as data from cities across the country show African American, Hispanic, low income and disabled students bearing their brunt (Oakes, 1999; Olson, 2001). Our own analysis of academic achievement measures[ix] for all Philadelphia eighth graders showed that if the promotion requirements planned for 2000-2001 had been implemented in June 2000, 60 percent of White[x] students would have been promoted to ninth grade without remediation, while only 31% of Black students, 35% of Hispanic students, 29% of low income students and less then 10% of students labeled learning disabled would have been eligible for high school.

These aggregate outcome data tell a discouraging story—that a student's background determines her fate. If this reform era is to generate educational equity for traditionally marginalized students, we need to ground outcome data within the context of specific local school practices. We need rich and detailed images of classroom interactions in which students, in partnership with their teachers, are working to higher standards. Moreover, we need to understand these interactions from students' perspectives so that we can draw lessons for developing reform policies that will engage young people meaningfully in their education and offer hope of truly effecting educational equity.

The data also point to the vast distance between policy's stated intentions—to enact higher standards for students—and practices on the ground. In that space, actors who include students, teachers, parents and others contest and negotiate the meaning and fairness of these new, increasing demands through their daily interactions. They experience policy individually without the benefit of a community where multiple viewpoints can be brought to bear on understanding and deliberating what happens in specific classrooms and the district at large. Yet, even as these actors are the ultimate shapers of reform, they rarely have opportunities develop and use collective knowledge to inform educational policies.

This is a critical time to document from students' perspectives how the current standards-based reforms are playing out in particular local contexts. The kind of layered understanding that links individual student's experiences and academic outcomes with specific reform policy requires longitudinal, multi-methods research. We believe that interpreting quantitative data about academic outcomes in conjunction

with richly textured pictures drawn from students' perspectives and experiences will help to explain the ways in which academic achievement and limitation are constructed in particular classrooms and schools. It will also help us understand how to influence the larger policy context of standards-based reforms and high stakes accountability, especially at the district level, but also on the national scene. Such research would provide a critical lens on current reform efforts and suggest the kinds of policies and classroom practices we need to interrupt negative patterns and increase positive outcomes for traditionally marginalized students (Firestone & Dawson, 1989; Erickson, 1986).

We are also convinced that this research must be conducted in close collaboration with those most affected by the new reform policies. Understanding these policies will not only require research on students, but research analysis with students, their teachers, and other school and district personnel. Our exploratory work this year encouraged us to imagine a structure (an interpretive community) and set of research practices that would productively bring local actors into research and reform. One critical lesson we took from our meetings with the interpretive community was that students and their parents must be part of this group in the future. At the same time, we came to understand how deeply practitioners desire opportunities to examine and speak back to these reforms in collaboration with colleagues, families and their students.

As we complete this essay, the Philadelphia Board of Education is weighing whether to implement the promotion policy it adopted in 1998 and the new U.S. President has proposed sweeping changes to Federal education policy that rely on the high standards reform policies the students in this study critiqued so thoroughly. That improving academic achievement, especially for urban, poor and minority students, is part of local and national debate is a positive development. This exploratory work suggests to us, however, that leaving students out of the conversation is a consistent and troublesome mistake. We hope to conduct a larger study that will demonstrate the benefits of the involvement of students and other local stakeholders in research and provide them with a voice for influencing the policies that most affect them.

V- References

Black, P. and William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139-148.

Cohen, D.K. (1990). A revolution in one classroom: The case of Ms. Oublier. *Educational Evaluation and Policy Analysis, 12*(3), 311-329.

Cohen, D.K. and Ball, D.L. (1999). *Instruction, capacity and improvement*. CPRE Research Report Series RR-43. Philadelphia: Consortium for Policy Research in Education.

Cohen, D.K. and Hill, H. (1998). *Instructional policy and classroom performance: The math reform in California*. CPRE Research Report Series RR-039. Philadelphia: Consortium for Policy Research in Education.

Corbett, D. & Wilson, B. (1995). Make a difference with, not for, students: A plea to researchers and reformers. *Educational Researcher, 24*(5), 12-17.

Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work.* San Francisco, California: Jossey-Bass Publishers.

Delpit, L. (1995). *Other people's children: Cultural conflict in the classroom*. New York: The New Press.

Education Trust. (2000). *Youth at the crossroads: Facing high school and beyond*. Prepared for the National Commission on the High School Senior Year. Author: Washington, DC.

Erickson, F. (1986). Qualitative methods in research on teaching. In M.C.Wittrock, (Ed.), *Handbook of Research on Teaching, Third Edition*. New York: MacMillan.

Firestone, W. and Dawson, J. (1989). Approaches to qualitative data analysis. In D. Fetterman (Ed.), *Qualitative Approaches to Evaluation in Education*. Praeger.

Fordham, S. and Ogbu, J. (1986). Black students' school success: Coping with the burden of 'acting white.' *The Urban Review, 18*(3), 1-31.

Fullan, M. with Stiegelbauer, S. (1991). *The New Meaning of Educational Change, Second Edition*. New York: Teachers College Press.

Hornbeck, D. and Ingram, G. (1998). *Policy 102 progress and plans: December 15, 1998*. School District of Philadelphia: Philadelphia.

Jencks, C. and Phillips, M. (Eds.) (1998). *The black-white test score gap*. Washington, DC: Brookings Institution Press

Jerald, C.D. (1998). The urban picture. In *Quality Counts 1998*. Washington, DC: Editorial Projects in Education

Johnston, R.C. and Viadero, D. (2000). Unmet promise: Raising minority achievement. *Education Week, 19* (27) p. 1, 18-19.

Knapp, M.S., Shields, P.M. and Turnbull, B.J. (1995). Academic challenge in high-poverty classrooms. *Phi Delta Kappan 76*(10), 770-776.

Kordalewski, J. (2000). *Standards in the Classroom: How teachers and students negotiate learning*. NY: Teachers College Press

Ladson-Billings, G. (1994). *The Dreamkeepers: Successful Teachers of African American Children.* San Francisco: Jossey-Bass Publishers.

Mitchell, R., Willis, M. and the Chicago Teachers Union Quest Center. (1995). *Learning in overdrive: Designing curriculum, instruction and assessment from standards*. Golden, Colorado: North American Press.

Oakes, J. (1999). Promotion or retention: Which one is social? *Harvard Education Letter, 15*(1): 8.

Olson, L. (2001, January 11). Finding the Right Mix. In Quality Count 2001 A Better Balance: Standards, tests and the tools to succeed. *Education Week, 20* (17), p. 12-20.

Ozga, J. (2000). Policy research in educational settings: Contested Terrain. *Philadelphia, Pennsylvania: Open University Press.*

School District of Philadelphia (1996). *Recommended content standards, benchmarks and performance examples.* Philadelphia.

School District of Philadelphia (1995). *Children Achieving action design.* Philadelphia: School District of Philadelphia.

Schwandt, T. (1994). Constructivist, interpretivist approaches to human inquiry. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research (pp. 118-137).* Thousand Oaks, CA: Sage Press.

Simon, E., Foley, E. and Passantino, C. (1998). *Making sense of standards: Implementation issues and the impact on teaching practice.* Children Achieving: Philadelphia's Education Reform Progress Report Series, 1996-97. Philadelphia: Children Achieving Challenge.

Spillane, J.P. (2000). *District leaders' perceptions of student learning (Report 0p-05).* Philadelphia: Consortium for Policy Research in Education.

U.S. Department of Education, Office of the Undersecretary, Planning & Evaluation Service. (2001). *High standards for all students: A report from the national assessment of Title I on progress and challenges since the 1994 reauthorization.*

Wasser, J. D. and Bresler, L. (1996). Working in the interpretive zone: Conceptualizing collaboration in qualitative research teams. *Educational Researcher, 25*(5): 5-15.

Wheelock, A.W. (1998). *Safe to be smart: Building a culture for standards-based reform in the middle grades.* Westerville, Ohio: National Middle School Association.

Wilson, B.L. and Corbett, H.D. (1999). *"No excuses:" The eighth grade year in six Philadelphia middle schools.* Philadelphia: Philadelphia Education Fund.

VI - Appendix A: Research methods

We conducted our exploratory study in two middle schools in Philadelphia. Tables 1 and 2 present important demographic and achievement information about each of the schools we studied, compared with Philadelphia middle schools in general.

Table 1. Characteristics of the two schools in the study, compared with Philadelphia middle schools in general.

School	Years of Principal Tenure	Number of Teachers	Number of Students	% Female	% African-American	% Asian	% Latino	% Native American	% White
Parks	15	16	244	52.9	66.4	3.7	1.0	0.8	28.3
Cobb	3	62	992	49.0	80.0	11.9	7.1	0.2	0.8
Middle Schools	Not available	Not available	876	50.2	67.5	3.6	13.4	0.2	15.2

Table 2. Achievement trends for the two schools in the study, compared with Philadelphia middle schools in general.

	Percent At or Above Basic SAT9 Math				Percent At or Above Basic SAT9 Reading			
School	1996	1997	1998	1999	1996	1997	1998	1999
Parks	20.3	28.4	37.0	32.4	64.9	80.3	74.1	76.1
Cobb	14.1	9.3	13.4	18.8	47.5	40.1	43.5	48.9
Middle Schools	15.7	18.5	25.2	24.4	43.3	50.5	55.0	58.5

Qualitative Data Collection and Analysis

Beginning in January 2000 through the end of the school year, we conducted six focus groups with students, interviewed teachers and principals (14), observed classrooms and faculty meetings (18), and reviewed regularly collected grade and behavior data kept in student files. The group of students we interviewed were representative of the racial/ethnic demographics in their schools, although not of the city overall. We interviewed three Asian-American students, six White students and 15 African-American students. Each researcher followed general protocols for interviews, focus groups and field note summaries. We obtained permission from parents to include their children in the study and to examine their student records.

Our work was also informed by an interpretive community of stakeholders—including teachers, researchers and central office personnel—that met twice during the research period. During our planning

phase we brought together individuals from a broad range of experiences, perspectives and positions within and outside of the school district, drawing upon Research for Action's strong history of working with teachers, parents, students, administrators and community members to define issues and questions, establish techniques for gathering information, and analyze data collectively. At the beginning of our research we invited a group of teachers and principals from a number of middle schools across the district to meet with us to surface emerging issues and questions. In fall 2000, we organized another meeting that brought together a group of teachers, principals, central office administrators and educators working in Philadelphia intermediary organizations to examine and interpret some of the data that we had collected.

Quantitative Data

We also worked with the School District of Philadelphia to obtain longitudinal demographic and academic outcome data for a large sample of 8th grade students in the District, in order to put our schools and the students we interviewed in wider context. We used these data to characterize our schools (see Tables 1 and 2 above), determine the proportion of students who would have met the promotion requirements, had they been implemented; to examine differences by race/ethnicity, class and disability and to compare teacher grades with standardized test scores. We found that:

- *Most students passed their courses, but far fewer scored at the required level on the SAT9.* While over ninety percent of middle school eighth graders for whom we have complete data passed all their major courses (with a "D" or better) in 1999-2000[xi], only 58% of the same group scored at or above Below Basic III on the reading and mathematics sections of the SAT9. These figures are significant because, at the time of our study, the School District of Philadelphia planned to incorporate these requirements for promotion to ninth grade.

- *Predictably, historically underserved students would bear the brunt of the negative consequences of the promotion policy.* If they had implemented those requirements in the 1999-2000 school year, only 37% of all Philadelphia middle schoolers would have met them without remediation. The outcomes are predictably worse for regularly underserved groups: only 31% of Black students, 35% of Hispanic students, 29% of low income students and less then 10% of students labeled learning disabled would have met the promotion requirements without remediation.

- *Students' level of success on teacher grades and standardized test scores were only minimally related.* The correlation between a student's grade in reading and their reading score on the SAT9 was 0.37; math grades and SAT9 scores had a correlation of 0.39 and in science the correlation was 0.36. These relationships are significant but small. Consider that the relationship between two standardized tests intending to measure the same construct is generally 0.7 or better.[xii]

Initially, we had also hoped to use these data to model student performance, identify atypical patterns of "success" and "failure" in the two middle schools and help us understand how classrooms and teachers shape outcomes for particular students. This analysis was hampered by several complications, which we will draw on to inform our larger study. The primary problem was that the school district did not release student test score data for the 8th graders we observed until December 2000, well into their ninth grade year. Our file was not ready until January, 2001. Under the parameters of this grant, we simply did not have time to complete an extensive regression analysis.

Exploratory work since then suggested that a deeper analysis would not have been fruitful due to several concerns:

- *Incomplete data.* When we received the data, much of the student information was incomplete. In some cases the reasons for missing data were typical: e.g. a student had missed his opportunities to take a test or had not provided the information we were seeking. In other cases, the missing data were the result of idiosyncratic data collection practices. For example, one of the schools where we conducted our study, Parks, uses an alternative grading scheme. Since their grades do not fit the typical A-F system, all of their grades are counted as missing in their pupil information file.

- *Unreported data:* The district did not store electronically students' grades on their multidisciplinary projects. Knowing who passed and who didn't in the full cohort of 8th graders is impossible without reviewing student transcript files.

- *Data accuracy:* There are some student background characteristics that are completed for every student, such as race and gender. These variables held fairly complete information for the 8th grade cohort, and, judging from our student files, the information was also accurate. In other cases however, only students having the characteristic defined by the variable are noted. That is, students who are enrolled in ESOL programs are marked by their ESOL level. For those students not in ESOL programs, the ESOL variable is simply left blank. There is no way to distinguish students who are not receiving ESOL services from those whose information is simply missing. The same is true for low income students, special education students, and drop-outs, etc. In such cases, data accuracy is suspect.

These issues suggest steps for us to take as we prepare for a full-fledged study of students' experiences with high standards reforms. The lead time for obtaining and cleaning quantitative data must be dramatically increased. We will also have to work more closely with pupil information services to minimize data limitations. We might also want to collect our own data more extensively.

[i] This research has shown that standards-based curriculum and instruction requires the consent and participation of teachers (Black & William, 1998; Cohen, 1990; Cohen & Ball, 1999; Cohen & Hill, 1998; Knapp, Shields, & Turnbull, 1995; Mitchell, et al., 1995; Simon, Foley & Passantino, 1998); that there are vast spaces between policy intentions and on the ground practices as teachers make complex instructional decisions that actively re-shape policy at the classroom level (Simon, Foley & Passantino, 1998; Cohen 1990); that policy makers have grossly underestimated the guidance and support necessary for teachers to make substantive changes in their practice (Christman, Corcoran, Foley Luhm, forthcoming); and, that policymakers continue to frame change as simply a matter of knowledge transmission, even when their ultimate policy goal is to promote social models of learning, such as constructivism (Spillane, 2000).

[ii] An exception to this is an interview study by Wilson Corbett (1999). This study, however, does not document students' experiences in actual classrooms. Another study documents students' experiences with exhibition projects. (Kordalewski 2000)

[iii] Wasser & Bresler's (1996) conceptualization of the "interpretive zone" in qualitative research teams was influential in our imagining an "interpretive community."

[iv] The School District of Philadelphia adopted a new promotion and graduation policy in 1998, and began to implement it in the 1999-2000 school year. The new policy eliminated social promotion and required significantly higher achievement from students at benchmark grades (4, 8 and 11) than the District ever had in the past. The policies varied slightly by grade level but generally, to move to the next grade or graduate all students had to pass all their major subjects, achieve a proficient score on standardized tests of reading and mathematics, take and complete more, and more difficult, courses and complete multidisciplinary and service learning projects. The requirements were also designed to be phased in, with requirements becoming gradually harder over time. For example, 8th graders in 2001 were expected to score at least Below Basic III on SAT9 reading and math tests; students who were 8th graders in 2004 had to score higher, at least the Basic level on these tests and so on. The Board adopted these requirements on the condition that funds to provide the supports necessary to help students reach these demanding requirements would be available. When, in June 2000, the State failed to provide additional funding to the School District, the Board postponed the implementation of the new requirements for 8th graders and graduating students. As of this writing, the Board and District leadership are still grappling with a decision about further implementation of the promotion and graduation policy

[v] The SAT-9 tests were also very important for schools. In 1996, the School District instituted an accountability system that held schools responsible for progress on the SAT-9s, students' promotion rates and student and teacher attendance rates.

[vi] See Appendix A for a more detailed account of the data we collected and for an analysis of our quantitative data set.

[vii] Most Philadelphia middle schools are divided into smaller units of 300-400 students and their teachers. Each SLC has a thematic focus.

[viii] The activities were shaped by Bloom's Taxonomy, a categorization of thinking skills ranked from the easiest to the higher order skills. The skills include demonstrating knowledge, comprehension, application, analysis, evaluation, and synthesis. In Bloom's taxonomy the last three categories are considered higher order skills. In the project there were five or six activity options under each Bloom category. Students were asked to complete one activity from each of the six ways of thinking.

[ix] See Appendix A for a summary of our conclusions from the quantitative analysis.

[x] These categories reflect school district's language for race and ethnicity.

[xi] It is interesting to note that the average grades given to students increased slightly in their 8th grade year, even though the trend throughout their middle school years showed average grades typically declining. This slight rise in grades might have stemmed from the fact that when teachers filled out their report cards they were under the impression that students who did not pass all four of their major courses would not be promoted to high school.

[xii] Correlation is measured on a scale of -1.0 to +1.0, with +1.0 signifying a perfect positive relationship, -1.0 representing a perfect negative relationship, and 0 representing no relationship.