

Accountability and Single-Sex Schooling: A Collision of Reform Agendas

Kathryn Herr

Montclair State University

Emily Arms

Loyola Marymount University

This ethnographic study documents how accountability measures skewed the implementation of gender equity reform at one California public middle school serving low-income students of color. In creating single-sex classes throughout the school, the Single Sex Academy (SSA) became the largest public experiment with single-sex schooling in the country, but pressure to raise its standardized test scores diverted the school away from the exploration and implementation of the gender reform. The chronicle of SSA is particularly relevant in light of (a) a recent call to relax Title IX standards and increase the numbers of public single-sex classes and schools, and (b) the provision of monies mandated by the No Child Left Behind Act of 2001 for single-sex classes and schools, along with the act's imposition of accountability standards and testing.

KEYWORDS: accountability, gender, single-sex schooling, standards.

As we celebrate the 50th anniversary of *Brown v. Board of Education*, a decision that desegregated schools by race, some African American historians of education have questioned the extent to which Black children benefited from that landmark case (Siddle Walker, 2000). A gifted Black child in a pre-desegregation school was known and nurtured by Black teachers and a Black community. The same child, bussed to a nearby White school,

KATHRYN HERR is a member of the faculty in the College of Education and Human Services at Montclair State University, Finley Hall 111, One Normal Avenue, Montclair, NJ 07043; e-mail herrk@mail.montclair.edu. Her research interests include issues of equity and education, youth studies, and practitioner and action research.

EMILY ARMS was a doctoral candidate in UCLA's Graduate School of Education and Information Studies when this study was completed. She is now an Assistant Professor at Loyola Marymount University, School of Education, University Hall, One LMU Drive, Suite 2600, Los Angeles, CA 90045-3772; e-mail: earms@lmu.edu. Her research interests include gender issues in K–12 and higher education, curriculum and instruction, and teacher education.

too often became just one more Black child whose special talents were invisible. Although few would want to return to a pre-*Brown* era, assumptions about segregation, whether by race, sex, class, or disability, must be tempered by research that provides data about the particular conditions under which segregating children in schools or among schools might lead to better social and educational outcomes. This article describes an experiment with segregating students by sex in a public middle school and explores how a mix of local conditions and a policy context stressing high-stakes accountability both promoted and distorted the promise of single-sex classrooms.

The Single Sex Academy (SSA) opened in 1999 to great fanfare. This California middle school, with all single-sex classes, became the largest public experiment in single-sex schooling in the country. The school is of particular interest in that it is a low-income, urban school serving students of color, a population not well served in many of our public schools. The first day of school was documented by local and national media, with television crews and newspaper journalists covering the story. “Boys and Girls in a Class Apart,” a headline announced in the *Chicago Tribune*, describing the advent of SSA as “an elaborate experiment to determine whether adolescent boys and girls learn better if separated from one another” (Haynes, 1999). The media documentation of the experiment continued throughout the school year.

We arrived to document the experiment as well. Although coming from different parts of the country and unknown to each other, we were both drawn to SSA because of its large-scale experiment with single-sex classes. Both qualitative researchers, we soon were meeting regularly, sharing insights derived from our data gathering; we found that our analyses cross-fertilized our individual research efforts. Initially, we were primarily interested in how students and teachers responded to the single-sex classes. Would they be able to create vital learning environments? Would they be able to address issues of gender equity? How did the teachers’ beliefs about gender affect their pedagogy? Our ethnographic approach allowed us to cast the net broadly. Although we had originally proposed to document what was presented specifically as a gender reform, we quickly saw that the multiple reforms being implemented in the school—reconstitution, high-stakes testing, standardized curricula and test preparation, and the single-sex, interdisciplinary classes—all had their own stories to tell. Although still interested in our original questions, we initiated a more open-ended search for the intended and unintended effects of multiple reforms on a single-sex experiment.

SSA was reconstituted¹ at what had been a neighborhood school. Almost from the outset, however, the faculty and staff experienced increasing pressure from the state and the school district to improve the school’s low ranking on the Academic Performance Index (API), primarily determined by student scores on standardized tests. In this climate, the success of SSA became equated with improving students’ performance on the Stanford 9 tests. Although that singular goal was accomplished—student test scores did increase in the first 2 years of the program—we would argue that it was not without consequences for curriculum, instruction, and possible gender reform.

This article chronicles how accountability measures affected and skewed the implementation of what was widely touted as being primarily a gender reform at one California public middle school. In the literature review we first provide some background information on the single-sex schooling debate in the United States and then discuss the current climate of high-stakes testing and accountability. We link both of these areas to current policy initiatives of the federal government. After a description of the school site and our research methodology, we present two strands drawn from our data analysis: (a) the impact of high-stakes testing on curriculum and on the roles of teachers; and (b) how the multiple reforms, simultaneously implemented, derailed the possibilities inherent in the single-sex experiment. In the "Implications" section of the article, we discuss the intersection of the local culture of the school with the larger policy contest in which it found itself; beyond this, we pose questions for ourselves and other educators as to how reforms can move from rhetoric to reality. We think the chronicle of SSA is particularly relevant given an ongoing call by the Bush administration to relax Title IX standards and to increase the numbers of public single-sex classes and schools. Simultaneously, the current administration, through the No Child Left Behind Act of 2001, has provided monies for experimenting with single-sex classes and schools while also imposing accountability standards and high-stakes testing.

Single-Sex Schooling

Educational research on single-sex classes and schooling is inconclusive and controversial. Although opinions are mixed, the Bush administration has signaled its intent to "promote single-sex schools" through the easing of a "rigid" interpretation of Title IX. As Mael (1998) points out, Title IX, perhaps unintentionally, has made public single-sex schooling in the United States virtually nonexistent. Provisions of Title IX state that no school receiving any federal funds shall "provide any course or otherwise carry out any of its education activity separately on the basis of sex" (Sax, 2003). Originally intended as a corrective to the sex discrimination and inequitable resource allocations found in coeducational institutions, it has also, ironically, curtailed some experiments with gender equity and reform such as public single-sex schooling.

In May of 2002, the U.S. Department of Education issued a "notice of intent to regulate," soliciting public comment on proposed changes to Title IX. The suggested changes directly challenged the athletic protections guaranteed under Title IX. At the same time the Department saw the possible loosening of the regulations as a way to encourage single-sex schooling (Davis, 2002). The Department asked the public to comment on whether the existing regulations "are appropriate or if they are unduly restrictive" ("ED to Seek Allowances," 2002). Seeing the athletic gains fostered by Title IX under attack, a fervent lobbying effort by over 100 groups (National Coalition for Women and Girls in Education, 2003) ensued. That effort resulted in an announcement in July 2003 that the Department of Education would not make any changes in Title IX (American Association of University Women

[AAUW], 2003). Although pleased with the verdict, the AAUW expressed concern that the Department of Education was considering changes to Title IX at all. It cautioned that “what will truly demonstrate the Bush administration’s commitment to Title IX and gender equity is the enforcement of Title IX policies” (p. 1). In March of 2004, the Bush administration reopened the issue of relaxing Title IX regulations, specifically to allow public schools to educate boys and girls separately (Knight, 2004). The period of public comment ended in late April 2004; it now remains for the Department of Education to sift through the public comments and make a final decision in the matter.

Possible tinkering with Title IX by the Bush administration has rekindled an old debate. Some, like the ACLU, have vehemently opposed any change in Title IX. The ACLU, in a letter of comment to the Department of Education, noted that “not only would the proposed changes in the regulations undermine Title IX and violate the Equal Protection Clause, but they would also perpetuate the inequalities associated with gender segregation” (“ACLU Single-Sex Notice of Intent Comments” 2002). The ACLU was joined by other organizations, such as the AAUW, the National Association of Women, and the Feminist Majority Foundation, which argued that segregation breeds inequality. Others, such as senators Hillary Clinton, D-N.Y.; Kay Bailey Hutchinson, R-Tex.; Barbara Mikulski, D-Md.; Edward M. Kennedy, D-Mass.; and Sue Collins, R-Maine, all have offered congressional endorsement of single-sex schools and single-sex classes (O’Keefe, 2002). In sum, what some see as increasing flexibility and options, others see as a dangerous erosion of civil rights.

Just as there is a lack of agreement in the policy arena regarding whether to further foster single-sex education, educational research on the subject presents some ambiguity and lack of agreement as to its benefits. There has been no national, comprehensive, controlled study of academic performance by U.S. students in public and private K–12 single-sex and coeducational schooling, in part because there are very few public single-sex classes or schools (Campbell & Sanders, 2002).² Although coeducation is virtually taken for granted in the United States, Riordan (2002) points out that it developed out of economic expediency³ rather than any visionary educational plan or research: “Historically, this mode of school organization was never subjected to systematic research” (p. 11).

Much of the current interest in reforms that revolve around gender dates back to work that emerged in the early 1990s, documenting that girls were being “shortchanged” in schools (AAUW, 1992) and that our public schools were “failing at fairness” (Sadker & Sadker, 1994) when it came to equitable treatment of boys and girls. These works documented a widespread bias against females in the current coeducational arrangements, which influenced educational achievements and possibilities. Riordan (2002) notes that, given these findings, placing the burden of proof on single-sex schools to show greater effectiveness than coeducational schools is a bit ironic. He suggests that coeducational schools should be asked to “demonstrate that they are at least as effective as single-sex schools in terms of achievement and gender

equity” (p. 11), essentially turning on its head the argument for the norm in education—coeducational schooling.

Just as an interest in single-sex education for girls was taking off, inner-city schools were experiencing a simultaneous movement focused on the education of minority boys, particularly African American boys (Salomone, 2000). All-boys’ classes and Afrocentric academies were seen as possible solutions in efforts to reverse the downward educational spiral of this population. In essence, the movement intertwined issues of race and gender, and some scholars noted that “just as coeducational classrooms have been shown to be hostile in many ways to female students, it has been argued that the coeducational environment is not optimal for African-American boys” (Singh, Vaught, & Mitchell, 1998, p. 158). Advocates of single-sex education, then, argue that this organizational structure provides a better learning environment not only for girls but also for at-risk urban minority students (Riordan, 1990). Drawing on research from private schools, Riordan (2002) indicates that, for historically or traditionally disadvantaged students (minorities and/or lower-class and working-class youth), the academic and developmental consequences of attending single-sex schools are significant.

Specifically, disadvantaged students in single-sex schools, compared to their counterparts in coeducational schools, have been shown to have higher achievement outcomes on standardized tests of mathematics, reading, science, and civics. They show higher levels of leadership behavior in school, do more homework, take a stronger course load, and have higher educational expectations. They also manifest higher levels of environmental control, more favorable attitudes toward school, and less sex-role stereotyping. They acknowledge that their schools have higher levels of discipline and order and, not surprisingly, they have a less satisfactory social life than students in coeducational schools. (p. 14)

But urban initiatives for developing all-male schools or academies, often specifically designed for students of color, have garnered vehement opposition. For example, the National Association for the Advancement of Colored People (NAACP) “insists upon retaining its historical opposition to ‘school segregation,’ ” (Hopkins, 1997, p. 8), arguing that segregation in any form could lead to forced resegregation. Legal challenges to all-male or all-female initiatives, regardless of whether they are linked to specific racial or cultural groups, have resulted in the closure of many of the single-sex initiatives.

In the United States, part of the rationale for single-sex schooling is the view that adolescents create a culture in school that is at odds with academic performance and achievement. Preoccupied with what Mael (1998) refers to as the “rating and dating” culture, students in the coeducational environment are thought to be unnecessarily distracted, concentrating on how they look rather than focusing on academics. Streitmatter (1999) notes that boys and girls distract each other and that this takes multiple forms, such as dressing

to impress the other sex, sexual harassment, and vying for teacher time and attention.

Although there are some studies of individual public schools such as the one presented here, educators in the United States have had to rely on studies conducted predominately in parochial schools (Riordan, 2002) or abroad, where single-sex school settings are much more common (Streitmatter, 1999). Because most of the research in the United States has been in private and parochial schools rather than in the public sector, selection bias confounds how the research results can be viewed (Mael, 1998). Given the current administration's policy thrust to develop more single-sex settings in public schools, attention must be paid to the emerging body of research in the United States drawn from public school settings.

Confounding the debate around a gendered reform in public schools and single-sex education are the measures employed to demonstrate the success of the experiment. If, as we observed earlier, success is equivalent to a rise in standardized test scores, what other aspects of the experiment are not being captured? And, as noted earlier, single-sex reforms do not usually take place in isolation from other massive reforms, such as high-stakes testing, that are currently popular in schools. This latter point is particularly important in light of Riordan's (2002) findings that, for example, the significant effects of single-sex schooling are small in comparison with the effects of the type of curriculum in a given school. When standardized tests are the indicators of success, and curriculum is crowded out by skill-and-drill types of pedagogy, the positive possibilities of single-sex schooling are compromised.

Constructing Gender in Schools and Classrooms

Although the AAUW (1998) reported that there is "no evidence that single-sex education is *better* than coeducation," it stressed at the same time that the long-term impact of single-sex education on girls or boys is unknown and that more research is needed. In recent years the AAUW has successfully placed the experience of girls in our nation's public schools on the national agenda. "In the field of education research, *How Schools Short-change Girls* sparked hundreds of new studies on girls in education, ensuring that girls became central to discussions in classrooms, schools, districts, states and across the nation" (AAUW, 1998, p. 2). In its most recent work on the topic, the AAUW (1998) observed that "public schools *are* making progress toward equitable treatment of boys and girls," but it called attention to the fact that shifting demographics indicating a more diverse student population "demand scrutiny of their impact on gender relations" (p. vii). The AAUW noted that the current scholarship needs to move beyond a discourse on gender equality—on whether boys and girls receive the same education—to grapple with gender equity that "addresses the needs of both girls and boys, rather than questioning whether each receives the same thing" (AAUW, 1998, p. 3).

Currently, there is a backlash against girl-centered research, such as that generated by the AAUW and others (Brown & Gilligan, 1992; Taylor, Gilligan, & Sullivan, 1995), which has moved issues of girls' development and education to front and center. The counter-question is, "What about the boys?" In the war of words, boys and girls are often compared to determine which group is currently more disadvantaged by educational practices. This comparison commonly uses standardized test scores, with the results collapsed into male and female categories for comparison.

A polarization has emerged that obscures equity issues and perhaps discourages needed research that would "address the needs of both boys and girls, rather than questioning whether each receives the same thing." Pollack (1998) points out that although groups of boys are the best performers in the schools by some measures, our focus on them as "academic superstars" skews our sense of the school experience for a broader range of boys. "There is a new gender gap with a predominant number of boys falling to the bottom of the heap" (Pollack, 1998, p. 234).

Whole groups of students continue to be disenfranchised in our current school and classroom arrangements, as attested to by the disproportionate number of poor and minority students who drop out of school. For example, Latinas and Latinos, the group of students increasing most rapidly in the United States, "continue to lag behind other ethnic minorities and Anglos on most measures of success, including educational attainment and its extension, economic well-being" (Ginorio & Huston, 2001, p. vii). "Thirty-one percent of Hispanic boys drop out, compared with 12.1% of black boys and 7.7% of white boys" (Canedy, 2001, p. 1). Boys of color who struggle to remain in school are disproportionately represented in the number of school expulsions, suspensions, and referrals to special education.

Clearly our current arrangements in school are neither equal nor equitable. There seems to be agreement that patterns of unequal support and attention are common in U.S. coeducational classrooms, from preschool through college (Mael, 1998; Sadker & Sadker, 1994), with females being "shortchanged" (AAUW, 1992). Other nuanced studies (Connell, 1996; Pollack, 1998) have begun to disaggregate boys as a block and are making the case that the intersection with race and social class stratifies the possibilities of success for boys in school, with minority boys particularly disenfranchised. Ferguson (2001) states that just as some children are tracked into prestigious professions, others are being prepared for prison. In her study of African American male students, Ferguson observes that it "became clear that school labeling practices and the exercise of rules operated as part of a hidden curriculum to marginalize and isolate black male youth in disciplinary spaces and brand them as criminally inclined" (p. 2).

Given the current state of affairs in our public schools, which clearly privileges some students while disenfranchising others, some see the standards movement as at least bringing these disparities to the public arena where they can be addressed. Work done in Texas by Skrla, Scheurich, and Johnson (2001) ties accountability and standardized testing to equity issues, arguing

that “accountability systems can play a key role in closing the achievement gap that historically has existed between the academic performance of White and middle-class children and that of children of color and children from low-income homes” (p. 228). They argue that by disaggregating data by race and social class, to see who is succeeding, deep inequities in traditional education settings can be revealed, forcing schools to deal with the demonstrated inequities. In the section below, we pick up the threads of the hotly contested argument that high-stakes testing and the standards movement can bring about equity in public schooling.

SSA Follows a National Trend: High-stakes Testing and Accountability

When the Bush administration took office, it quickly moved to make its mark in the educational policy arena. Building on what some (Fuller & Johnson, 2001; Koschoreck, 2001; Sclafani, 2001; Skrla & Scheurich, 2001) have heralded as the great success in Texas of using accountability measures and standardized testing to foster school reform, what was once state policy quickly became the national approach to educational reform. When Bush signed the No Child Left Behind Act in January 2002, it contained several important provisions that have relevance for the study reported here. For one, it put into place a national educational accountability approach, with high-stakes testing the most common measure, and it started the time clock for states to devise accountability structures to demonstrate that “no child would be left behind.” At the same time, embedded in the act were provisions and monies to promote experimentation with single-sex classes and schools; \$3 million in appropriations for single-sex schools and classrooms were part of the bill’s “innovative programs” (Schemo, 2002).

High-stakes accountability policy has moved to the forefront of a national debate. Proponents of accountability argue that disaggregation of data from high-stakes testing can reveal who is succeeding and who is not in our schools, forcing schools to deal with the demonstrated inequities. But the possibility of a social equity agenda linked to high-stakes accountability has been critiqued by various authors (Anderson, 2001; McNeil & Valenzuela, 2000; Parker, 2001). It is worth quoting Parker (2001) at length on this critique:

Whites will indeed tolerate any policies that have a harmful impact on racial minorities, but if the tables were turned and Whites suffered under these same policies, then they would work actively to change them. Thus, Whites have no problem accepting and even advocating for higher assessment standards or grade retention policies, even though, according to House (1999) much research shows how harmful these policies are on students, just as long as they mostly affect minority students in urban schools. However, once the policies affect Whites, particularly those in the suburbs, then Whites will react to change these policies. . . . If the test’s covert intent is to boost minority urban education achievement, then it is a targeted policy that basi-

cally lets White students “off the accountability hook,” and this potential differential impact may need to be explored not only from a psychometric standpoint but also from a critique of whiteness and White privilege. (pp. 314–315)

House (1999, cited in Parker, 2001) makes the argument that policies with enormous racial implications are often “disguised as nonracial. . . . This disguise enables the policies to seem fair and democratic, even when the policies have racial overtones” (p. 2). House links educational practices such as standardized testing, retention, and ability grouping, among others, as ways to ensure that students of color are saddled with an inferior education.

Work by McNeil and Valenzuela (2000) illustrates how the implementation of one of the policies House questions, that of standardized testing, resulted in an inferior education for many students of color. They found that the particular standardized test used in Texas, the TAAS, created many harmful effects on teaching and learning, such as reduced class time for high-level instruction, increased money spent on test-prep materials that diverted scarce instructional dollars away from high-quality instructional resources, increased minority student dropout rates, increased numbers of students labeled Special Education (to keep them from being tested), and reduced quality and quantity of the subject matter.

All states have developed testing programs, and most have aligned tests, curriculum, and state standards (Cimbricz, 2002). In many states, the tests have high stakes—rewards or sanctions—attached to them. High-stakes decisions with serious consequences for students, such as promotion or graduation, are in place. In addition, the work of teachers and administrators is also “measured” by the results of their students’ standardized test scores. Whole schools risk being placed on probation or run by state governments until the scores improve; private companies can be brought in to take over the leadership of the school, displacing decision making by local professionals who are seen as having failed. The schools in particular jeopardy are often those in urban, poor neighborhoods such as the one in which SSA is located.

Although accountability programs involving standardized tests have been put in place in virtually all states, there is very little research indicating that education or learning improves with accountability, particularly when it is equated with scores on standardized tests. Schools are increasingly accountable to people outside the field of education, such as politicians and businessmen; the concept of accountability has been a political one from its inception (Huebner & Hauser, 1999), and its political aspect has escalated with the Bush administration and legislation such as No Child Left Behind.

Anderson (2001) argues that as accountability is increasingly equated with high-stakes testing, the focus on the test

narrows and impoverishes the curriculum, creating opportunity costs that lead to the elimination of more in-depth analysis of critical social

issues. As the state takes a stronger role in curriculum and testing . . . standardized testing can become a more sophisticated technology of control. (p. 323)

Hand-in-hand with high-stakes testing is the creation of an environment where the discretion previously exercised by teaching professionals has given way to intensified surveillance on the part of administrators and district offices. Standardized curricula are mandated; administrators are held accountable for implementing them and therefore bring this sense of surveillance into the classroom.

At the heart of the debate over high-stakes testing is the question whether it can be leveraged to help raise the performance of schools whose student bodies are primarily poor, non-White, and urban, or whether it helps to perpetuate an inferior education. Interestingly, one of the arguments for single-sex classes is that they may be of particular benefit to those currently most disenfranchised in traditional education settings—the same poor, non-White, urban students (Riordan, 2002; Singh, Vaught & Mitchell, 1998). What we have then, in this study, is numerous reforms being implemented for the benefit of students who have experienced failing schools in the past, with the hope of increasing educational benefits for them through various reforms. As we will describe, the discrete parts do not necessarily add up to a better whole.

The School Site

SSA is a public, urban middle school comprised of more than 1,100 girls and boys in Grades 6–8. With nearly 95% minority students, the racial composition of the student body is 44% Latino (many of whom are English Language Learners), 28% African American, 22% Asian American (primarily of Cambodian descent), and 6% White. Nearly 80% of these students come from lower-socioeconomic-status households and receive free or reduced-price lunches. Although all students in the district are eligible to apply to the school, the overwhelming majority of the student body comes from the surrounding neighborhood.

The academic core classes are offered in interdisciplinary blocks (100 minutes of English/history and 100 minutes of math/science), and each academic teacher has a block of girls and a block of boys. This arrangement, where each teacher has a class of boys and a class of girls, was done with an eye toward Title IX and to comply with the federal legislation. For us as researchers, the unique arrangement of the school's classes offered a natural environment in which to study differences in instructional practices in classrooms that separate students by sex.

It is important to note that like many urban schools with poor minority populations, SSA was under tremendous pressure from the district and state to improve its standardized test scores. In the process of becoming a single-sex magnet school, SSA was “reconstituted” by the district. The main reason cited by district personnel for reconstitution was the fact that the school ranked

second-to-last among 14 middle schools in the district in terms of Stanford 9 test scores. In California, where this study took place, assessment of public schools was based on a single measure: students' Stanford 9 test scores. Reconstitution, a drastic maneuver, meant that the district replaced all administrative positions at the school, and all teachers were required to reapply to teach at the school. In the process, SSA lost many of its veteran teachers. Replacing them were teachers new to the school and the neighborhood, many of whom lacked experience and/or were on emergency credentials.

In addition to an overwhelmingly new teaching staff, new curricula were instituted as well. These included a remedial reading program and a test preparation program. An emphasis by the administration on improving test scores led many of the newer teachers (as well as the veterans) to rely heavily on teacher-centered lessons—such as timed practice tests, skill-and-drill, or lectures—that required more disciplinary control of the students. Instructional format came to be aligned with test taking.

While the stated goal of SSA was to tailor instruction at the school to meet the needs of the new all-boy and all-girl classes, the single-sex discourse remained at a level of “minimizing distractions.” Despite the description offered by the school district that “there has been extensive training for all staff in issues of gender equality,” during our 2 years at the school we saw only one instance of professional development that related explicitly to the single-sex arrangement. That even took place at the first faculty meeting of the new school: A film featuring William Pollack's work with boys was shown, and a guest speaker who had experience working in an all-girls' school spoke with the faculty.

Against this backdrop, teachers at SSA began the first year struggling with numerous, sometimes conflicting reforms, while venturing into uncharted territory together with their students. But coupled with this was also a sense of excitement as the school garnered national publicity from its gender experiment and was the subject of numerous newspaper and television stories following its progress as the largest public school gender experiment in the nation's history.

Research Methods

As stated earlier, we were initially interested in how the single-sex experiment would affect teaching and learning. We wondered how teachers would grapple with the new gender reform and how their own beliefs about gender might affect their pedagogy. Beyond this, we were curious to see if the single-sex experiment could further gender equity. However, as we came to realize how many reforms were being implemented simultaneously at the school, we decided, in addition, to explore the effects of multiple reforms on the single-sex experiment.

The data in this article are part of a larger data set that includes teacher, administrator, and student interviews; classroom observations; open-ended student surveys; and document analysis. This article draws from teacher inter-

views, classroom observations, and, to a lesser extent, student interviews. Our other data also inform the knowledge base for the article.

For purposes of this article, we draw primarily on repeated teacher interviews and classrooms observations of 18 core teachers and 36 classrooms. In the school there were a total of 32 core teachers (those who taught the interdisciplinary blocks in major subject areas) and 18 other faculty who taught physical education, electives, and special education. Of the 18 participating teachers, we selected 6 from each grade level (sixth, seventh, and eighth). Nine were English/history core teachers and 9 were math/science core teachers. Our focal teachers varied in age from 24 to 54 and ranged in teaching experience from a first-year novice to a 20-year veteran who had won the district's Distinguished Teaching Award in 2000. As in many urban schools, the student population was very diverse, but the teaching staff was predominately White. Therefore most of our focal teachers were White, although 2 Latino and 4 Asian American teachers participated.

We interviewed students in each classroom that we observed, as well as more informally on the playground and at lunch; this provided background information for us, although we did not include much of it in this article. The typical number of students in a class was 35; of those, more than half were interviewed in each class, in part because there were so many eager volunteers. The students interviewed were Latino, Asian American, African American, and Anglo-American. Students were interviewed in pairs because research indicates that they are more comfortable with that approach (Fine, 1988).

Participant Observation

As mentioned above, the primary methods of data collection were participant observation and in-depth interviews. We used these data collection methods for the purpose of building the richest description of the single-sex classrooms at SSA. We observed each teacher's all-female and all-male classes back-to-back (for a total of 3 hours each time), between four and eight times. On other days, we followed the students as they went from one core teacher to the next (also three hours); we did this multiple times. We wrote field-notes of these observations, recording in a regular and systematic way what we observed and learned from the daily routines of the classes. The field-notes of classroom observations yielded the emerging themes and patterns that became a core of our data and fed into our interview questions.

In-Depth Interviews

In addition to participant observation, we used in-depth, semi-structured interviews of the participating teachers. Teachers were interviewed between two and four times. Students were generally interviewed once. Interviews are valuable in uncovering "participants' meanings for events and behaviors and for generating typology of cultural classification schemes" (Marshall & Rossman, 1989, p. 112). All interviews were tape-recorded and transcribed. Interviews

were used to discover the participants' views, particularly on gender and the reforms at the school site. For the teachers the interview topics ranged from questions about their professional experience and staff development to the various reforms instituted at the school and their impact on curriculum and instruction. The interviews generally lasted between 45 and 60 minutes because they were typically done during teacher preparation periods or at lunch. Each teacher was interviewed at least twice, once before we observed her or his classroom and once after the observation.

Students were released for interviews during homeroom, lunch, and class time. Their interviews were typically 30 to 40 minutes long. We interviewed students after observing in their classrooms multiple times, in the hope that we would become familiar faces to them. Interview questions revolved around their experiences in the single-sex classrooms as compared with their previous coeducational experience.

Analysis

We employed a grounded theory approach to our data (Glaser & Strauss, 1967). Although as researchers we were drawn to studying the gender reform at SSA, grounded theory requires that the data tell the story of what is most salient at a given research site. In preliminary coding of teacher interviews and classroom observations, we became aware of the impact of test preparation and curriculum standardization on the single-sex experiment. Using the constant comparative method of returning to the data again and again, certain themes emerged that had little to do with gender, directly. Both our observations and interviews vividly and repeatedly illustrated the all-consuming focus on standardized test preparation at SSA. Thus a story began to emerge of how the gender reform was eclipsed by an extreme emphasis on improving students' standardized test scores at all costs and by the impact of that focus on curriculum, instruction, and students. The data presented in this article are not extreme examples or exceptions; rather, they are the best examples among multiple instances. We have chosen only representative excerpts of coded interviews and classroom observations to illustrate our main themes.

Trustworthiness of the Study

We were privileged to be involved with SSA over a sustained period of time. One of us was in the school for a year, the other for 2 years. Our prolonged engagement and persistent observation (Lincoln & Guba, 1985) allowed us to see how the story unfolded over time. In addition, we were able to triangulate our data because we drew on multiple sources from varying perspectives, as well as multiple data-gathering methodologies, for our information. Beyond this, we employed member checking throughout the study. We were fortunate in that the principal asked us to present our findings to the faculty during faculty meetings. In addition, we often engaged in conversations with faculty about what we were discovering and how we were making meaning; this allowed us to get their responses, verification, and continued thinking

on the themes we were discovering. At the end of the study, faculty and administration were given a written “executive summary” of findings, to which they were asked to react and respond.

Findings

In the following sections, we unpack the multiple intersecting reforms being simultaneously implemented in the school and explain how they created challenges for teachers. In exploring the impact of high-stakes testing, we look first at the move to mandated, standardized curricula. The result was a predominately teacher-centered, less authentic, and less creative style of teaching. Teachers found it more difficult to cultivate student learning in engaging ways. We then explore the de-skilling and control of teachers. Issues of controls on teachers are shown to trickle down to students, in that more relational teaching receded and the teaching task became one of controlling bored students. We link the issue of control to the single-sex classroom arrangements. Uninterrogated gender beliefs are shown to reproduce racist and sexist classroom environments for students.

The Impact of High-Stakes Testing

The pressure to raise the school’s Academic Performance Index (API) permeated the air at SSA. The single-sex reform would be considered a success if the scores went up. On her desk the principal had a small, carved number 16, the number of points the school needed to gain on the Stanford 9 tests to meet its prescribed goal for the year. The resources allocated for professional development related to the single-sex experiment were much more limited than those related to test score improvement. For example, every classroom had copies of *Test Best*, a Stanford 9 test-prep curriculum. By the second year of the reform, the district required SSA to adopt the packaged reading program *Soar to Success*. Published by Houghton Mifflin, *Soar to Success* is a “reading intervention program” for students in Grades 3–8 who are reading “significantly below grade level.” Thus, although the tests took just a few days in the spring, the push to raise the scores infused much of what teachers did or did not do in their classrooms all year long.

The Impact of Testing on Curriculum

From our fieldnotes we offer the following example of a teacher’s struggle to continue to offer a creative curriculum in the face of the curriculum designed for test preparation:

Mr. C. is perched on a tall stool, a pencil behind his ear. Two boys are distributing copies of *Phantom of the Opera*. For the past few days they have been discussing the various roles in a play or a movie—the narrator, the director, etc. Today they will continue reading through the play, with the boys cast in all the parts, male and female. Mr. C. tells the boys, “There’s a romantic exchange between Raoul and Chris-

time; there are no girls in here so you need to rise to the occasion.” He taps Pablo to be Christine, saying, “Pablo’s an actor; he can handle it.” Pablo nods in agreement. Mr. C. continues going around the room, assigning parts. Meanwhile, a boy turns to Pablo and whispers, “They kiss” to which Pablo replies, “It’s only a play.”

The boys begin to read the play. The Phantom emits a scary “Christine!” “Good job,” Mr. C. murmurs. There’s an extended exchange between Raoul and Christine; Pablo as Christine sings some of his lines in a high-pitched voice. They get to the love scene; the rest of the boys are quiet and appear absorbed. They reach the end of the segment they are doing for this class period.

Mr. C. puts on the tape of the opera. The boys are listening, a few moving slightly to the music. A boy is following in the script with his finger, quietly mouthing the words along with the tape of Christine’s singing. Mr. C. coaches them on what to listen for: “Really focus on the emotion going on, particularly when the Phantom enters the scene.” Pages of the script turn together. I’m struck by how rare it is to get to listen to something beautiful in school.

Following the opera the boys return to a written assignment they had started the day before. I overhear Mr. C. say to Pablo, “Very good job today—not only reading that part but singing it. Taking leaps and bounds toward manhood.” The bell rings and the boys file out. I tell Mr. C. how much I enjoyed the class; he thanks me but anxiously replies that tomorrow he has to get back to *Test Best*, that he’s worried they’ve lost too much time doing *Phantom*. (Fieldnotes, 2/28/00, pp. 5–7)

The initial space for teachers to design lessons that were creative and more hands-on was giving way to prepackaged curricula designed to facilitate student success on the Stanford 9 tests. In the first year of the reform, *Test Best* was initially used once a week during the 20-minute homeroom period; but as the testing time drew nearer, the use of *Test Best* in the classroom increased to several times a week. By the second year of the reform, in addition to mandating what teachers taught, teachers were told *how often* to use the materials.

Teachers expressed concern with the intrusion on curriculum and the teaching of larger concepts; they felt instead that the emphasis had become more skill-and-drill oriented, with concepts taught out of context.

I feel that I’m spending way too much time on test preparation. Not teaching to the test, but taking skills out of context and, like, “You got to know this; you got to know this; you got to know this!” And at the expense of, like I mentioned, history. There’s just not time for the whole picture kind of thing. (Interview, 3/19/01, p. 8)

Because of the emphasis on English and computational skills on the standardized tests, other subjects were often ignored or given short shrift. This was especially hard on history teachers, who because of interdisciplinary blocks were teaching English and history, and science teachers, who taught math and science. The blocked classes provided 50 minutes on each

subject, or the option of integrating the two areas throughout the 100-minute class; however, as testing dates approached, many teachers spent the entire 100-minute block on English or math. One experienced history/English teacher wistfully recounted the influence of mandated test prep on his history curriculum:

But yes, it's [test prep] adversely affecting the teaching in the sense that I'm doing a lot less history. I used to do TCI, which is that history slide lecture program, and it's got a lot of interactive games and interactive programs of projects. And I haven't been able to do those. I did them at the beginning of the year, but once the memos started coming out about getting kids ready for their tests, I had to kind of stop that. (Interview, 3/3/01, p. 17)

Not all teachers let the one tested subject take over their 100 minute core class, although those who persisted in teaching both subjects did so with some trepidation at deviating from what was becoming the norm at SSA. As one science teacher explained, who taught both math and science,

I taught science all year. I know some of my colleagues didn't, but I did. I didn't think that was right. If I am given two curricula to teach, I am going to teach them. I generally taught both. . . . I rarely took my science time and taught math like I know a lot of my colleagues did. And maybe that's a mistake in the sense that my kids will do worse on the math SAT9 [Stanford 9]. But I don't think it's a mistake; I really like science and they *need* science. (Fieldnotes, 6/13/00, p. 1)

In contrast to the novels and plays that English teachers taught during the first year of our research, the texts used in *Soar to Success* during the second year were often nonfiction picture books on famous people or types of animals. Only in the "high" English classes (which were tracked by student scores on the algebra portion of the standardized tests) did we ever observe students reading real literature texts. For example, in one class where the students were divided into five reading-level groups, only the highest group was reading a fiction book that could be considered literature; the other four groups read nonfiction books with pictures.

De-Skilling and Controlling Teachers

With success of the reform being equated with increasing the school's API ranking, we saw an ongoing transformation of approaches to teaching and learning. Teachers lamented that using a prepackaged test-preparation curriculum meant giving up their own lesson planning. They struggled with the realization that they were losing the purposes and joy of teaching. They had the sense that what they were doing was not "real" teaching, although they still struggled at times to hold on to their original vision.

Math is what I have to teach for the exam. All my lesson plans are pretty much written for me for that. I don't get any creativity. Though

I do try to add my own with real-life links. I try linking math with science; I can't just teach them separately. I think it will benefit the students in the long run, but I have to look at the state test as the final assessment, which pushes to teaching to the test. Which I don't think is the best way to teach a kid. Every Friday we take a little quiz that's based on the state exam. It's timed. So that's what we do: We prepare them to take a timed test. (Interview, 2/7/00, p. 13)

Many teachers felt that, for most of the school year, anything not directly tested on the Stanford 9 was not a priority to teach. At the end of the second year of the test prep curriculum, one veteran teacher summed up her feeling that basically test prep is not good teaching:

As far as the testing affecting the success of the student, what it did was it took time away from good instruction, I think. It was too much skill-and-drill instead of good instruction. (Interview, 3/19/01, p. 7)

The change from good instruction to skill-and-drill was not lost on the students. One history/English teacher observed that when he did teach "authentically," as he had done "before test prep," the students noticed the difference:

Yesterday, when I was lecturing on how feudalism began, one of the students said, "Mr. G., you used to talk like this, you used to teach us like this in the beginning of the year. I like it when you teach this way; how come you can't talk like this everyday?" I said, "Gosh, I would love to, but they won't let me." But it made me feel good, you know, history is so interesting. (Fieldnotes, 3/03/01, p. 18)

Both teachers and students noticed and lamented the shift in the classroom environment when test preparation took precedence over other teaching styles. Teachers disappeared behind scripted curricula and students resisted the loss of more creative, authentic learning experiences through disengaged or "bad" behavior.

Previous research has described in detail the deleterious impact of "controlling" reforms on teachers. As McNeil (1986) writes, "bureaucratic controls on teaching [are] not vague influences, but rather very concrete and visible transformations of course content and classroom interaction" (p. 11). Controls from above, such as standardizing the curriculum, contribute to the "de-skilling" of teachers through the taking away of creative spaces where teachers are full participants in cultivating their students' learning. But as McNeil (1986) points out, teachers also participate in their own de-skilling, exerting controlling practices over their own students, recreating in their own classrooms the hierarchical relationships district administration and standardization have imposed on them. At SSA, interactions with students were more scripted; classes were more rote, and it was increasingly difficult to "control" the bored students. Ironically, the single-sex classes, implemented without

any substantive professional development that might have helped the teachers to interrogate their own gendered beliefs, contributed to the growing sense of needing to control the students. The intersection of these reforms is explored in the next section.

The Impact of Single-Sex Classes

In the creation of SSA, the rationale was very much along the lines of “minimizing distractions” (Mael, 1998); this became a mantra of sorts, repeated often by faculty and students, and easily the most commonly offered explanation for the new configuration of the school. When scratched beneath the surface, however, the distraction explanation revealed little thought or teacher preparation in terms of a broader agenda of school reform that addressed issues of gender bias or equity. Educators were essentially left to do the best they could in responding to the single-sex configuration. Without ongoing staff development or conversations regarding gender, there was nothing to interrupt teachers’ gendered assumptions and ideologies. There was no impetus to help them problematize their own understandings or standpoints, and no catalyst to realizing new pedagogical possibilities that might foster equity. As one teacher observed,

There really hasn’t been much guidance. Um, straight from the principal’s mouth, “We took the distractions out of the classroom and that’s the bottom line.” That’s what it’s all about, just separating boys and girls, so they would not distract each other. Anything beyond that is just from my own interest. (Interview, 11/5/99, p. 4)

The faculty seemed generally aware of reports such as the AAUW’s (1992) *Shortchanging Girls*, and generally knew about the rationale for intervening in the middle school years when girls are at a “crossroads” in terms of self-esteem and achievement. As one teacher commented, “All of this was intended originally to allow them [the girls] to be free of the male influence” (Interview, 12/8/99, p. 2). The faculty were less clear on how the experiment might benefit the boys. Almost immediately, the discourse surrounding gender reform in the school went in two distinct directions: It was widely believed that the single-sex experiment was “working for the girls,” but there was much less enthusiasm for the all-male groups.

Single-Sex All-Male Classes

Although there was, of course, a continuum of teaching and classroom cultures in the all-male classes, much of what we saw revolved around control issues. It was widely believed that teachers “couldn’t let up for a minute” in the all-male classrooms or things would careen wildly out of control. This was attributed, in part, to the absence of girls in the classroom and the view that boys couldn’t control themselves. One eighth-grade teacher observed:

The boys are free to be just as good or bad as they want to be, . . . but in many cases they are going to wind up being just as bad as they want to be. Because before, let's say that theoretically, you have half the class in the disruptive male influence, now you've got the entire class as the disruptive male influence. . . . So, for teachers, they have 35 kids potentially cutting up as opposed to 15, or whatever they had before. (Interview, 1/6/00, p. 4)

Many boys, too, agreed that the all-male classes were struggling without the girls in the classroom.

Like in elementary, when they used to put a boy between girls, I think the girls help, . . . 'cause there's, like, all boys now and there's, like, more of—everybody's rowdy. There's not somebody to keep us down a bit. (Interview, seventh-grade boy, 2/4/00, p. 3)

I think boys do some things they wouldn't do when they are around girls. They [the girls] keep us in check. (Interview, sixth-grade boy, 11/16/99, p. 2)

Actual classroom observations of the boys' classes revealed a range from “hyper-controlled” classrooms, where the teacher insisted on absolute obedience, to classrooms in wild disarray with none of the boys apparently paying attention. What the ends of the continuum held in common was an implicit, and often explicitly stated, belief by both male and female teachers that boys could not control themselves—that they were to be controlled.

After a day in which the all-male classrooms were totally “out of control,” teachers often resorted to what McNeil (1986) has termed “defensive teaching”: Teachers anticipated student resistance and simplified course content, demanding little from the students and reducing knowledge to a list of facts to memorize for tests. The following example is drawn from notes on an observation conducted in an eighth-grade, all-male classroom:

Boys come spilling into the room, some leaping over desks on their way in; Ms. R. insists they go back out and enter the room again. Once they are in the room she tells them that today there will be no warnings, that the first infraction will earn them a call home. She switches on the overhead where the word respect is written; the rest of the page is covered up. “Respect,” she announces, “copy this definition in your notebooks.” She unveils the rest of the overhead and reads aloud: “Respect is doing what the person in charge of you tells you to do.” Boys are opening notebooks and looking for writing utensils. Ms. R. continues, “The first thing you've got to learn here is to respect the teaching going on and the learning of fellow students.” She tells them they are not to question assignments and suggests that outside of class they confront those who are “pulling the class down—talk to them outside of class, tell them what they're doing.” The room is now somewhat quiet and the boys are writing.

Ms. R. has written their assignment on the board—to copy pages of the science text into their notebooks. “The first thing I’ve asked you to do is a no-brainer. Hopefully something will stick in your head.” Ms. R. went on to say that the second assignment will be a “bit more;” it involves looking up a word. The room is quiet. Some boys are not writing at all; they have their heads down on their desks, eyes closed. There is an occasional noise in the classroom but overall it is “peaceful.” The boys continue to copy the science book for the first hour of class. (Fieldnotes, 10/21/99, pp. 1–3)

In an interview following this observation, the teacher tearfully confessed, “I am so afraid of them, so afraid” (Interview, 10/21/99, p. 8).

Analogous to Ferguson’s (2001) account of “troublemaking” students transformed into “troubled” children who need treatment, whole classrooms of boys became “troubled” classes in need of intervention. The issues presented in these classrooms were not so much pathologized as criminalized. Ferguson (2001) observes that, as children’s behavior is refracted through the lens of criminal perpetrator, it is “adultified”:

By this I mean their transgressions are made to take on a sinister, intentional, fully conscious tone that is stripped of any element of childish naïvete. The discourse of childhood as an unfolding, developmental stage in the life cycle is displaced in this mode of framing school trouble. (p. 83)

Key to understanding these “troubled classrooms” of boys is how gender intersected with race and social class, presenting teachers with a collage of images and beliefs to interrogate. In the United States there is a societal discourse, fanned by media representations, that urban boys of color are a menace to be feared and controlled; school environments are no exception to this gendered stereotype (Welch, Price, & Yankey, 2002).

Coupled with the belief that urban students, particularly the boys, need to be controlled was the environment that teachers found themselves in as professionals; they had little autonomy, little time or space to be creative and nurture learning. As mentioned earlier, the daily scripted tasks took time away from teachers’ relationship building with their students, because there was a sense that off-script interactions were “off task” and kept students from practicing skills needed for the tests. Many of their own teaching tasks were highly scripted and, ironically, while disillusioned about their own scripted tasks, many imposed similar tasks on their students, particularly the boys. Some of this was the nature of the scripted curricula and practice tests, but this was melded with an uninterrogated view of urban boys of color as fearsome. The learning atmosphere was a stifling one, in many cases, with whole blocks of classes being spent on rote learning—either as a punishment (that is, a form of defensive teaching) or from fear. Teachers feared their male students, or they feared that their classes would not perform well enough on the standardized tests unless inordinate amounts of time were devoted to skill-and-drill.

To echo and expand upon a point made by Datnow, Hubbard, and Conchas (2001), we saw “evidence of how teachers’ ideologies about gender”—and here we would add race and social class—“and the interactions between teachers and students in the classroom mediated the implementation of single-gender public schooling” (p. 23). Teachers’ beliefs and ideologies about gender, as well as about race and social class, played out in varying ways in their classrooms; but the larger issue is that there was nothing in place structurally to help teachers interrogate these beliefs and consider how they intersected with their pedagogy.

There are no guarantees that simply separating the sexes creates an equitable learning environment or one that interrupts stereotypical gender and racial arrangements. The same educational arrangements (single-sex schooling) that Riordan (2002) has indicated can offer particular promise to students who are marginalized and disadvantaged can also perpetuate the discouragement and alienation that is so familiar to these students in school. In the case of SSA, the opportunity to think about pedagogy in the single-sex classroom was displaced by issues of control and the pressure to increase standardized test scores.

What About the Girls?

When we asked to observe classes, teachers were often eager to show us their all-female classes and reticent about the boys’ classes. There was a solid sense in the school that this experiment, “designed for the girls,” was working for them. As one teacher commented, “All of this was intended originally to allow them [the girls] to be free of the male influence. To either not worry about being too smart or not worry about being too dumb” (Interview, 2/8/00, p. 2).

Teachers reported, and observations confirmed, that the girls’ classes were much less challenging in terms of overt discipline or classroom “control” issues. A number of teachers stated that once they got through their male classes for the day, they could relax in the all-female rooms, confident that the worst was behind them. Many teachers, both male and female, commented on how difficult it was to switch gears from what was required of them in the all-male classes—typically a more authoritarian approach—to the requirements of all-girls’ classes, where issues of control could recede somewhat to the background. Because a productive classroom climate was more easily established in the all-female classrooms, we expected to see issues of pedagogy move to the forefront as teachers could think more creatively about fostering a gender-equitable climate, at least for the girls.

But what we saw clearly was the tension to teach to the test and to cover the standardized curriculum, even at the expense of a teachable moment in terms of gender. We offer the following example in this regard, drawn from a history class where a teacher referred to Harriet Beecher Stowe’s *Uncle Tom’s Cabin* in a Civil War lesson. At first the students did not realize that Stowe was a woman; when they did, hands shot up in the air:

"She was a woman?" asks one girl incredulously.

"Yes, Harriet Beecher Stowe was a woman," replies the teacher.
[There is general excitement and talking among the girls.]

"A woman could write a big important book way back then?"
asks another student.

"Yes, yes," says the teacher. "Okay we need to move on so we'll
have time for the grammar quiz at the end of the period."

"Awwwe," several students seem disappointed. (Fieldnotes,
3/18/02, p. 2)

These teachable moments came and went, with teachers feeling pressure to stay on track and spend their classroom time "on task." We saw these moments when gender lessons could be drawn explicitly to the forefront of the class, but we rarely saw teachers capitalizing on them to "disturb" students' gendered beliefs. As teachers struggled to implement the mandated test-prep and packaged reading program, authentic teaching got lost along the way. With their freedom to design the best curriculum to meet the needs of their students diminished, formerly energized teachers succumbed to following the scripted curriculum. A seventh-grade girl observed:

Ms. R., she acted like it was hard for her to wake up; she was, like,
"Oh, I'm just tired of you kids." Just because we're young doesn't
mean we don't deserve respect too. (Interview, 4/4/00, p. 7)

In their study of "how America's schools cheat girls," Sadker and Sadker (1994) documented that in much of girls' school experience, they are "short-circuited by adults"; the authors explain that, whereas teachers are likely to show boys how to do something that challenges them, they are more likely to do a challenging task for girls. Ultimately the girls hang back, having lost faith in their skills, and let boys take the lead in the classroom, to the point where the boys dominate classroom interactions. Part of the hope offered by single-sex classrooms was that in all-female spaces, girls would more fully emerge into confidence and positive leadership skills.

As we stated earlier, there are no guarantees that simply separating the sexes creates an equitable learning environment or one that interrupts stereotypical gender and racial arrangements (Datnow, Hubbard, & Conchas, 2001). It was not unusual in the all-girl classrooms to see them essentially reproduce what had transpired in a typical coeducational classroom, but this time with some girls now taking the lead roles: A few girls would dominate, taking much of the "air time" in class, garnering a lot of teacher attention, both positive and negative. Just as a few boys in coeducational classrooms become the "bad boys" and take up a lot of teacher time, a few girls in each classroom emerged in these roles. As one eighth-grade girl commented, "I think the behavior is worse because the girls don't feel shy any more if they get in trouble" (Interview, 2/7/00, p. 3). Another group of eighth-grade girls agreed with this:

They [the girls] act even worse than when they were in classes with
the boys [chatter of agreement on this point with girls talking over

each other]. They think because they're not in class with boys they can show off more. (Interview, 5/4/00, p. 5)

Many other girls would sit quietly, not interfering or interrupting either. These girls were in familiar territory, learned during their coeducational classroom days.

In their studies of classroom interactions, the Sadkers (1994) documented “the silence of girls from grade school through graduate school” (p. 90); they also found that twice as many boys as girls argue with teachers when they think they are right. For many of the teachers who had already spent part of their day in classrooms where boys challenged virtually any classroom activity, it made a certain kind of sense to leave the girls undisturbed; silence had to be a welcome relief of sorts. We rarely saw the girls’ classrooms “out of control”; but, then, neither did we see evidence of innovative pedagogy that might encourage voice or leadership on the part of girls. Because many of the girls were “good schoolgirls”—that is, appeared to listen in class, did their assignments, and handed in homework—teachers were satisfied that the reform was “working for the girls.”

This brings us full circle in asking how we look at success in single-sex classrooms and whether maintaining the status quo with girls is enough. If single-sex classrooms are to reach their potential, attributes of a successful classroom must be made explicit, linked to gender equity goals. In a climate where standardized test scores are the one valued measure, SSA was a “success.” But as we have shown here, if the lens on its success includes interrupting stereotypical gendered lenses on the part of teachers and students, it has a long way to go.

Implications

It is important to remember that SSA was a school that had been reconstituted, an increasingly common intervention for schools with a poor showing in the high-stakes testing arena (Malen, Croninger, Muncey, & Redmond-Jones, 2002). Reconstitution destroys the culture of a school (Sergiovanni, 2000), making it particularly vulnerable “what culture the mania of high stakes testing brings” (p. 1).

Sergiovanni (2000), drawing loosely on the work of Habermas (1987), describes the ideally symbiotic relationship between the *systemsworld* and the *lifeworld* of schools. The systemsworld is the “world of instrumentalities” (p. 4), the policies, management systems, rules, and schedules that provide a framework for students and teachers to engage in the practices of teaching and learning. In the current context of high-stakes testing, where schools are issued report cards published in local newspapers, the systemsworld dominates school practices and cultures.

The lifeworld is the culture of the school, reflective of the values, norms, and beliefs of a school community, that determine the social interactions between students, teachers, and administrators; it “provides the foundation for the development of social, intellectual, and other forms of human capital

that contribute, in turn, to the development of cultural capital which then further enriches the lifeworld itself" (Sergiovanni, 2000, p. 5). A school that has been reconstituted has to rebuild its school culture and is doing so at a time when much is being imposed upon it (scripted curriculum, high-stakes testing, etc.) from beyond the school itself.

Ideally, the systemworld and lifeworld function in balance, creating schools and classrooms in which system efficiency is in harmony with the lifeworld where teaching and learning take place. But as Sergiovanni (2000) points out, when the systemworld dominates the lifeworld, as it did at SSA, this unbalance destroys the fabric of the school culture, resulting in isolation, alienation, and loss of professional efficacy. This imbalance characterizes the current standards movement in education, resulting in what Habermas (1987) calls the colonization of the lifeworld of schools.

We have brought to the local level the implementation of policies crafted at the state and national levels under the banner of an improved educational experience for those students most at risk: low-income, urban students of color. Although some scholars (Fuller & Johnson, 2001; Skrla & Scheurich, 2001) have expounded on the potential of standardized curricula and high-stakes testing to provide more equitable educational opportunities for traditionally underserved populations, we did not see evidence of this at SSA. What we found was a school on the move with the rising test scores to prove it, while the lived experience in classrooms was one of a narrowed, reduced curriculum, less meaningful lessons and learning, and less authentic teaching. The changes were reminiscent of what Kohn (2000) describes as "raising the scores, ruining the schools." SSA was declared a success on the basis of its increased standardized test scores, despite other data such as ours that would call that "success" into question.

Previous research (Riordan, 2002; Singh, Vaught, & Mitchell, 1998) has indicated that single-sex classes may benefit those urban students of color, particularly males but females, too, who typically have not been well-served in coeducational classrooms. The danger in portraying SSA as a success is that it short-circuited the potential reform that might truly benefit the students involved, and instead offered a cheap substitute: achievement based on skill-and-drill. SSA can be *portrayed* as a successful innovation when in reality it is a repackaging of the same short-shrift education that typically is offered to this population of students. The myth that students are receiving a better educational experience, based on raised test scores, ensures that SSA will continue on its current path and that the possible benefits of single-sex classes will remain unexamined and unachieved. A reform that could be truly innovative remains unimplemented but is acclaimed as a success. We are in a period of smoke-and-mirrors education, perpetuating educational inequities while indicating otherwise through our standardized testing and curricula.

The change in the quality and quantity of curriculum at SSA is all the more disturbing in light of research that has documented the negative effects of test preparation on poor, urban youth. With their higher risk for dropping out of school, underrepresented students particularly need relevant, engaging

lessons. To merely substitute test prep for authentic curricula is discriminatory and racist. As McNeil (2000) writes:

When standardization becomes institutionalized, and student testing comes to be used for monitoring “accountability” throughout a state’s educational system, the negative effects fall most heavily on the poorest children, minority children whose entire school experiences come to be dominated by an attempt to raise their (historically low) test scores at any cost. (p. 17)

Overall, SSA stands out as a school with multiple and perhaps well-intentioned reforms that were derailed by standardization, accountability, and failure to examine gender and racial beliefs. We would argue that any comprehensive school reform is severely limited when standardization and accountability measures are allowed to be the overarching innovation (see, e.g., Gunzenhauser & Gerstl-Pepin, 2002, for a discussion of how arts integration was affected through accountability measures in North Carolina).

Much of the research on single-sex schooling indicates that single-sex classes alone cannot address the structural barriers to gender equity. As indicated earlier, there are no guarantees that simply separating the sexes creates an equitable learning environment or even one that would be considered “better” in terms of interrupting stereotypical gender relations (Datnow, Hubbard, & Conchas, 2001; Zwerling, 2001). Without a careful and consistent interrogation of racialized, social-class-based, and gendered relations in a classroom and school, these innovative schools are in danger of perpetuating rather than interrupting stereotypes that damage students’ educational and future possibilities. This is particularly devastating given the preliminary evidence that single-sex classes and schools might indeed offer possibilities to underserved populations.

In the United States we are currently far from Sergioivanni’s (2000) suggestion of the ideally symbiotic relationship between the systemsworld and lifeworld of schools. In initiating No Child Left Behind, the Federal government has intruded further into what was previously the domain of state and local governments. This intrusion has led to a dominant role for federal policies and management, essentially unbalancing the more ideal symbiotic relationship. But we see reassertions at the state level, calling into question and even resisting some of the tenets of No Child Left Behind. As a result, the Bush administration is easing some of the provisions of the law (Dillon, 2004). The current unbalance, then, between the systemsworlds and lifeworlds of public schooling is fluid, affected by the agency of various stakeholders.

Currently there is a healthy debate among scholars about whether high-stakes testing can help create equitable schooling for students who typically are disenfranchised. Although there is some reason to think that single-sex classes and schools can be part of a move toward equity, scholars have had less opportunity to research the issue in public school settings in the United States. Through its repeated review of Title IX and allocation of funding for

single-sex experiments, the Bush administration has signaled its intent to move forward with more public single-sex classrooms and schools. These experiments will be important ones for gender scholars to document.

Our own research argues against simplistic reforms and measures to assess progress on equity issues. As we have seen in the case of SSA, a school can be portrayed as a successful single-sex experiment on the basis of gains in test scores without having addressed, as an institution, the issue of equity. As new single-sex experiments are developed, and as researchers study them, the challenge will be to go beyond these simplistic measures to more authentic assessments to move schools toward genuine gender equity. Ironically, we would suggest the same in terms of assessing student gains: to go beyond simplistic measures to more authentic assessments to move schools toward genuine equity. We echo Golnick and Chin (1994) in envisioning education where

Students' educational and vocational options are not limited by gender, age, ethnicity, native language, religion, class or disability. Educators have the responsibility to help students contribute to and benefit from our democratic society. Diversity is used to develop effective instruction strategies for students in the classroom. In addition, . . . education should help students and teachers think critically about institutionalized racism, classism, and sexism. Ideally, educators will begin to develop individual and group strategies for overcoming the debilitating effects of these societal scourges. (p. v)

Notes

The authors would like to thank Gary L. Anderson, Betsy Noll, Lee Anne Bell, and anonymous reviewers for comments on earlier drafts of this article.

¹Reconstitution typically means that administrators and teachers in a given school are removed and replaced. "Advanced as a bold, 'results-based' action that will 'turn around' troubled schools, reconstitution is a prevalent but understudied strategy" (Malen, Croninger, Muncey, & Redmond-Jones, 2002, p. 113).

²The Department of Education has just launched what is believed to be the first comprehensive study of public single-sex schools in the United States.

³Riordan (2002) points out that coeducation came into being because of financial constraints. "Historically, mixed-sex schools were economically more efficient" (p. 10). The result is that in the United States boys and girls typically attended the same public schools.

References

- American Association of University Women. (1992). *How schools shortchange girls: The AAUW report*. Washington, DC: American Association of University Women Educational Foundation.
- American Association of University Women. (1998). *Separated by sex: A critical look at single sex education for girls*. Washington, DC: American Association of University Women Educational Foundation.
- American Association of University Women. (2003). *Huge victory on Title IX*. Washington, DC: American Association of University Women Educational Foundation. Retrieved on July 12, 2003, from <http://www.aauw.org>
- American Civil Liberties Union. (2002). ACLU Single-Sex Notice of Intent: Comments to the Department of Education. Available from ACLU Online Archives at <http://www.aclu.org>

- Anderson, G. L. (2001). Promoting educational equity in a period of growing social inequity: The silent contradictions of Texas reform discourse. *Education and Urban Society*, 33(3), 320–332.
- Brown, L. M., & Gilligan, C. (1992). *Meeting at the crossroads: Women's psychology and girls' development*. Cambridge, MA: Harvard University Press.
- Campbell, P. B., & Sanders, J. (2002). Challenging the system: Assumptions and data behind the push for single-sex schooling. In A. Datnow & L. Hubbard (Eds.), *Gender in policy and practice: Perspectives on single-sex and coeducational schooling* (pp. 31–46). New York: Routledge Falmer.
- Canedy, D. (2001, March 25). Often conflicted, Hispanic girls are dropping out at high rate. *New York Times*, sec. 1, p. 1.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practices. *Education and Policy Analysis Archives*, 10(1). Available at <http://epaa.asu.edu/epaa/v10n2.html>
- Connell, R. W. (1996). Teaching the boys: New research on masculinity and gender strategies for schools. *Teachers College Record*, 98(2), 206–235.
- Datnow, A., Hubbard, L., & Conchas, G. (2001). How context mediates policy: The implementation of single-gender public schooling in California. *Teachers College Record*, 103(2), 184–206.
- Davis, M. (2002). Department aims to promote single-sex schools. *Education Weekly*, 21(36), 24–25.
- Dillon, S. (2004, March 14). U.S. set to ease some provisions of school law. *New York Times*, sec. 1, pp. 1, 21.
- ED to seek allowances for single-sex class schools. (2002). *What Works in Teaching and Learning*, 34(10), 2. Retrieved 5/25/02 from <http://proquest.umi.com/pqdweb>
- National Association for Single Sex Public Education (2003). *Federal notice regarding single sex education in public schools*. Poolesville, MD: Author. Retrieved on March 7, 2003, from <http://www.singlesexschools.org/legal.html>
- Ferguson, A. A. (2001). *Bad boys: Public schools in the making of Black masculinity*. Ann Arbor: University of Michigan Press.
- Fine, G. (1988). *Knowing children*. Newbury Park, CA: Sage.
- Fuller, E. J., & Johnson, J. F., Jr. (2001). Can state accountability systems drive improvements in school performance for children of color and children from low-income homes? *Education and Urban Society*, 33(3), 260–283.
- Ginorio, A., & Huston, M. (2001). *Si, se puede! Yes, we can: Latinas in school*. Washington, DC: American Association of University Women Educational Foundation.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Gollnick, D. M., & Chin, P. C. (1994). *Multicultural education in a pluralistic society*. New York: Merrill.
- Gunzenhauser, M. G., & Gerstl-Pepin, C. (2002). Guest editors' introduction: The shifting context of accountability in North Carolina and the implications for arts-based reform. *Educational Foundations*, 16(2), 3–14.
- Habermas, J. (1987). *The theory of communicative action: Vol. 2. Lifeworld and system: A critique of functionalist reason*. Boston: Beacon Press.
- Haynes, V. D. (1999, September 30). Boys and girls in a class apart. *Chicago Tribune Internet Edition*. Retrieved on September 30, 1999, at <http://www.chicago.tribune.com>
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Hopkins, R. (1997). *Educating Black males: Critical lessons in schooling, community, and power*. Albany: State University of New York Press.

- Knight, H. (2004, March 4). Plan to let schools divide sexes: Bush proposal to pay for all-girl or all-boy class. *San Francisco Chronicle*. Available at www.sfgate.com
- Kohn, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.
- Koschoreck, J. W. (2001). Accountability and educational equity in the transformation of an urban district. *Education and Urban Society*, 33(3), 284–304.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Mael, F. (1998). Single-sex and coeducational schooling: Relationships to socioemotional and academic development. *Review of Educational Research*, 68(2), 101–129.
- Malen, B., Croninger, R., Muncey, D., & Redmond-Jones, D. (2002). Reconstituting schools: “Testing” the “theory of action.” *Education Evaluation and Policy Analysis*, 24(2), 113–132.
- Marshall, C., & Rossman, G. B. (1989). *Designing qualitative research*. Newbury Park, CA: Sage.
- McNeil, L. M. (1986). *Contradictions of control: School structure and school knowledge*. New York: Routledge & Kegan Paul.
- McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.
- McNeil, L. M., & Valenzuela, A. (2000). The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric. In G. Orfield & M. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high stakes testing in public education*. New York: Century Foundation.
- National Coalition for Women and Girls in Education. (2003). Women and girls to President Bush: “Hands off Title IX!” Retrieved on January 29, 2003, from <http://www.ncwge.org>
- Oakes, J., Quartz, K. H., Ryan, S., & Lipton, M. (2000). *Becoming good American schools: The struggle for civic virtue in educational reform*. San Francisco: Jossey-Bass.
- O’Keefe, M. (2002, December 8). Bush initiative rekindles debate over gender-specific education. *The Seattle Times*. Retrieved on March 16, 2003, from <http://seattletimes.nusource.com>
- Parker, L. (2001). Statewide assessment triggers urban school reform: But how high the stakes for urban minorities? *Education and Urban Society*, 33(3), 313–319.
- Pollack, W. (1998). *Rescuing our sons from the myths of boyhood*. New York: Random House.
- Riordan, C. (1990). *Boys and girls in school: Together or separate?* New York: Teachers College Press.
- Riordan, C. (2002). What do we now know about the effects of single-sex schools in the private sector? Implications for public schools. In A. Datnow & L. Hubbard (Eds.), *Gender in public policy and practice: Perspectives on single-sex and coeducational schooling* (pp. 10–30). New York: Routledge Falmer.
- Sadker, D., & Sadker, M. P. (1994) *Failing at fairness: How our schools cheat girls*. New York: Simon and Schuster.
- Sadker, M. P., & Sadker, D. (1986). Sexism in the classroom: From grade school to graduate school. *Phi Delta Kappan*, 67, 512–515.
- Salomone, R. (2002). The legality of single-sex education in the United States: Sometimes “equal” means “different.” In A. Datnow & L. Hubbard (Eds.), *Gender in public policy and practice: Perspectives on single-sex and coeducational schooling* (pp. 47–72). New York: Routledge Falmer.
- Sax, L. (2003). What is the legal status of single-sex public education in the United States? Poolesville, MD: National Association for Single Sex Public Education. Retrieved on May 29, 2003, <http://www.singlesexschools.org/legal.html>

- Schemo, D. J. (2002, May 9). White House proposed new view of education law to encourage single-sex schools. *New York Times*. Retrieved May 26, 2002, from <http://www.nytimes.com>
- Sclafani, S. (2001). Using an aligned system to make real progress for Texas students. *Education and Urban Society*, 33(3), 305–312.
- Siddle Walker, V. (2000). Valued segregated schools for African American children in the South, 1935–1969: A review of common themes and characteristics. *Review of Educational Research*, 70(3), 253–285.
- Singh, K., Vaught, C., & Mitchell, E. (1998). Single-sex classes and academic achievement for all children. *Journal of Negro Education*, 67(2), 101–129.
- Skrla, L., & Scheurich, J. J. (2001). Displacing deficit thinking in school district leadership. *Education and Urban Society*, 33(3), 235–259.
- Skrla, L., Scheurich, J. J., & Johnson, J. F. (2001). Toward a new consensus on high academic achievement for all children. *Education and Urban Society*, 33(3), 227–234.
- Streitmatter, J. L. (1999). *For girls only: Making a case for single-sex schooling*. Albany: State University of New York Press.
- Taylor, J. M., Gilligan, C., & Sullivan, A. M. (1995). *Between voice and silence: Women and girls, race and relationship*. Cambridge, MA: Harvard University Press.
- Tyack, D., & Cuban, L. (1995). *Tinkering toward Utopia: A century of public school reform*. Cambridge, MA: Harvard University Press.
- Welch, M., Price, E. A., & Yankey, N. (2002). Moral panic over youth violence: Wilding and the manufacture of menace in the media. *Youth and Society*, 34(1), 3–30.
- Zwerling, E. (2001, June 3). California study: Single sex schools no cure-all, *Women's eNews*. Available at www.womensenews.org

Manuscript received December 8, 2003

Revision received February 6, 2004

Accepted April 11, 2004