Tracking In The Era of High-Stakes State Accountability Reform: Case Studies Of Classroom Instruction In North Carolina

MAIKA WATANABE
San Francisco State University

Background/Context: Considerable controversy surrounds the issue of whether high-stakes statewide accountability programs have led to more equitable educational opportunities for all students. Some researchers suggest that these programs have focused attention on improving the achievement of students of color from low socioeconomic backgrounds. Critics of accountability programs, however, raise concerns that high-stakes standardized tests have lowered the quality of instruction in inner-city schools, where students of color from lower socioeconomic backgrounds have traditionally not tested well, as these students do more out-of-context math and reading drill in preparation for the exams. A central question remains: What is the nature of curriculum and instruction for different groups of students in the new school reform context of high-stakes, statewide accountability programs, and what are the implications for equity?

Purpose of the Study: This account of classroom instruction delves into the instructional opportunities afforded students across different academic tracks under North Carolina’s accountability program. The author focuses on the nature of classroom instruction for students in the “regular” classes, which are disproportionately populated by students of color from low socioeconomic backgrounds, with that of their peers in “academically gifted” classes and considers the implications for equity in this new policy context. While previous tracking literature suggests that teacher expectations of students, material available to students and skills emphasized in instruction differ across tracks, the research community has conducted little empirical research to support or refute claims on differential learning opportunities for different groups of students in the new context of high-stakes statewide accountability programs.

Research Design: This research used an ethnographic case study approach with theoretical sampling to select participants.
Data Collection/Analysis: This research was based on the following data: 1) 68 hours of classroom observations in two focal language arts teachers’ classrooms, 2) six interviews with each of the two focal teachers, 3) student work from two focal teachers’ classrooms and 4) interviews with eleven other teachers teaching at other schools.

Conclusions/Recommendations: The analysis reveals five key differences in instruction across tracks that favored students in the “academically gifted” classes. The article discusses the extent to which the high-stakes accountability policies influenced teachers to make bifurcated curricular decisions for their students across tracks. The author argues that despite calls to close the achievement gap through high-stakes accountability programs, the gap will persist unless policymakers and educators consider ways in which school organization can perpetuate or equalize instructional opportunities for students.

INTRODUCTION

Considerable controversy surrounds the issue of whether high-stakes statewide accountability programs have led to more equitable educational opportunities for all students. Some researchers suggest that these programs have focused attention on improving the achievement of students of color from low socioeconomic backgrounds (Skrla, Scheurich, Johnson & Koschoreck, 2000). Critics of accountability programs, however, raise concerns that high-stakes standardized tests have lowered the quality of instruction in inner-city schools, where students of color from lower socioeconomic backgrounds have traditionally not tested well, as these students do more out-of-context, math and reading drill in preparation for the exams (Delpit, 2003; McNeil, 2000; McNeil & Valenzuela, 2001). A central question remains: What is the nature of curriculum and instruction for different groups of students in the new school reform context of high-stakes, statewide accountability programs, and what are the implications for equity?

This account of classroom instruction delves into the instructional opportunities afforded students across different academic tracks under North Carolina’s accountability program. I focus on the nature of classroom instruction for students in the “regular” classes, which are disproportionately populated by students of color from low socioeconomic backgrounds, with that of their peers in “academically gifted” classes and consider the implications for equity in this new policy context. While previous tracking literature suggests that teacher expectations of students, material available to students and skills emphasized in instruction differ across tracks (Oakes, 1985; Page, 1989, 1991), the research community has conducted little empirical research to support or refute claims on
differential learning opportunities for different groups of students in the new context of high-stakes statewide accountability programs.

BACKGROUND

In 1996, North Carolina implemented a high-stakes accountability program that rewards all teachers at a school with incentive pay if their students’ standardized test scores meet or exceed state expectations for growth. The key components of North Carolina’s high-stakes accountability program are as follows: 1) standardized testing aligned with state standards, 2) school classifications based on school-wide growth in achievement and percent of students performing at grade level, and 3) rewards, including up to $1500 cash bonuses for individual teachers, and negative sanctions attached to school classifications. Two priorities set the tone for the state policy context. From the accountability program’s inception, policymakers sought “to improve achievement ...and to offer aggressive assistance, as well as attractive incentives, to local schools” (North Carolina Department of Public Instruction, 1999, p. 4). Since 1999, policymakers highlighted a second priority, to close the achievement gap between students of color and their White peers, by establishing a Closing the Achievement Gap section within the Division of School Improvement in the North Carolina Department of Public Instruction. State policymakers currently emphasize the potential of the accountability program to reach both goals.

During the 2001–2002 school year when I collected data for this research, the North Carolina Commission on Raising Achievement and Closing Gaps reported on the state’s progress to reach both goals. While the percentage of students performing at grade level had improved for all ethnic groups, a thirty-point gap remained between White and Black students performing at grade level statewide on multiple-choice reading and math tests in grades 3–8. The achievement gap had decreased by only 4.3% in the five years since policymakers implemented the high-stakes accountability program in 1996 (North Carolina Commission on Raising Achievement and Closing Gaps, 2001). Since 2001, however, the gap has decreased, but a 21.8% gap still remained in 2005 (North Carolina Department of Public Instruction, 2005).

This article discusses one of the complex reasons why the state’s emphasis on “accountability for equity” by itself will not be able to challenge the persistent patterns of the achievement gap. Certainly, factors such as poverty, racism, family issues, school structures and teacher expectations contribute to the achievement gap in interconnected ways.
This article focuses on some of the school-based factors using qualitative data to illuminate the classroom practices of well-intentioned, hard-working teachers, who negotiate teaching in the context of the high-stakes accountability reform.

SIGNIFICANCE OF THE STUDY

Research on the influences of high-stakes statewide accountability programs on classroom instruction is particularly crucial as policymakers and educators across the country respond to the No Child Left Behind (NCLB) legislation that requires states to implement such programs. North Carolina is an especially interesting context to study for two reasons. First, North Carolina was one of the first states to pioneer high-stakes statewide accountability programs in the early to mid-1990s. North Carolina’s program has thus been a sustained, prominent reform for almost a decade. Secondly, North Carolina, along with Texas, has received national acclaim for progress in raising student achievement (Grissmer & Flanagan, 1998; Linn, 2003). In fact, the NCLB legislation “was modeled, in part on North Carolina and Texas’ accountability programs” as a result of the progress students in these states made (North Carolina Department of Public Instruction, Communications Division, 2002, p. 1). Therefore, a close examination of classroom instruction under North Carolina’s well-established accountability program can inform the efforts of other states now responding to NCLB legislation.

To date, research on the influence of high-stakes statewide accountability programs on classroom instruction suggests that teachers self-report aligning their classroom instruction to testing demands, emphasizing tested subjects and specific tested components (Hillocks, 2002; Jones et al., 1999; Kelley, 1999; Kelley & Protsik, 1997; McNeil, 2000; Siskin, 2001). However, few empirical studies have focused on the nature of classroom instruction for different groups of students attending the same school in the new context of high-stakes statewide accountability programs. In addition, few researchers incorporate sustained classroom observations in their research designs in this new policy context.

This piece extends previous research on the influence of high-stakes accountability policies on classroom instruction in four ways. First and foremost, my research unites two bodies of literature, tracking literature and literature on high-stakes accountability programs, which are ordinarily kept separate. I study classroom instruction across tracks in the new context of high-stakes accountability programs that the state says is a
Tracking In the Era of High-Stakes State Accountability Reform

major component in its efforts to close the achievement gap. Secondly, this article adds to the small body of literature on high-stakes statewide accountability programs that makes classroom observations a central data point (McNeil, 2000). This methodological choice allows data triangulation between what teachers self-report in interviews and actual classroom practice. Thirdly, this research focuses on a non-award winning site in contrast to previous research on award-winning school sites (e.g., Kelley & Protsik, 1997), which allows an opportunity to investigate complex patterns at a school site that is similar to most schools in the state. Finally, I collected data six years after implementation of the statewide accountability program in North Carolina, at a point when the statewide accountability program was firmly established, in contrast to previous research conducted in the early years of the North Carolina accountability program (Jones et al., 1999).

This article also further extends previous research on tracking. Researchers such as Oakes (1985) and Page (1989, 1991) focused the scope of their research on the classroom level, comparing instruction across tracks. Their contributions brought greater familiarity with classroom-level mechanisms that contribute to the effect of tracking on students’ schooling experiences, but the state policy contexts of the classrooms were largely absent from their analyses. My research builds on this prior work on tracking by asking the following research questions: In the new policy context of high-stakes statewide accountability programs, where state policymakers have directed attention to not only raise student achievement but also to close the achievement gap, what is the nature of classroom instruction for students across tracks? How has the nature of classroom instruction across tracks changed, if at all, in this new policy context from patterns documented in literature prior to high-stakes testing?

This article enables us as policymakers, researchers, teacher educators and practitioners to further consider the equity implications of accountability policies for different populations of students. It also offers insight into the complex factors that contribute to the achievement gap and focuses attention on alterable school-based factors to foster more equitable instructional opportunities for all students.

METHODS

This study takes its point of departure from the perspective that public schooling is contested territory, an arena where students, teachers, administrators, parents, policymakers and community members interact in complex ways to produce shifts that move toward equality or social
reproduction (Carnoy & Levin, 1985). For example, research literature suggests that teachers can greatly alter the intent of policy reforms at the implementation stage by fully or partially rejecting (or accepting) policy reforms based on their knowledge, beliefs about their students and interpretation of classroom events (Ball & Bowe, 1992; Cohen, 1996; McLaughlin, 1987). Thus, focusing a research lens on how actors inside the “blackbox” of schooling construct meaning of their multiple embedded contexts and how they enact classroom practices is powerfully informative, both in terms of the intended and unintended consequences of policy reform.

An ethnographic case study approach guided my research because I wanted to get an in-depth picture of the complex patterns unfolding at the classroom level and of teacher perspectives on their curricular decisions over the course of a school year, patterns that most likely would not have emerged without the depth of case study research. I used theoretical sampling to select participants in the study. Theoretical sampling is a process by which researchers select particular cases in specific contexts in order to generate theory (Glaser & Strauss, 1967). My case study research is thus not designed to make generalizations about a broader population of teachers. However, the consistent pattern of findings across classrooms and schools in my research indicate the need for further investigation into the classroom-level influences of high-stakes accountability programs for different groups of students using a larger, randomly selected sample.

PARTICIPANT SELECTION

Grade Level and Subject Area

In order to examine the role of testing in teachers’ decisions about what they taught across tracks, I first narrowed the realm of possible subject areas to research a tested subject, language arts. Seventh grade was the focus of my research since seventh graders took writing tests in addition to reading and math standardized, multiple-choice end-of-grade tests. The extra writing test at the seventh grade level, where students had to write a clarification or point-of-view essay, presented an added challenge for seventh grade teachers under the statewide accountability program (Testing Section, NCDPI, 1999).

Focal School

I selected Sequoia Middle School, a school similar to most North
Carolina middle schools—a school that had a comprehensive mission, operated on a traditional school year calendar, and where, like 54.2% of the middle schools in the state, teachers had not won reward money as part of the state’s accountability program in 2001 (North Carolina Department of Public Instruction, 2001).

Focal Teachers

Because of the intensive demands of ethnographic research and my solo efforts, I elected to concentrate on two teachers. The main criterion I used to select the focal teachers, Ms. Jones and Ms. Royster, out of the four seventh grade language arts teachers at Sequoia Middle School, was their years of experience. I ensured that these focal teachers had more than three years of experience so that I could concentrate on the influence of the state accountability policy rather than the demands that beginning teachers face in their first three years of teaching (Huberman, 1989; Veenman, 1984).

Ms. Jones, a White teacher who graduated from a state teachers’ college with a BA degree in English, was in her thirtieth year of teaching seventh grade language arts at Sequoia Middle School. Ms. Royster, a Black teacher who received a BA in English and minors in education and sociology from a traditionally Black liberal arts college, was in her fourth year of teaching.

Teachers at Other Schools

In addition, I recruited eleven seventh grade language arts teachers at four other middle schools in the school district with different school reward histories (River, Cornwallis, Rose Valley and Jefferson) to participate in interviews at the end of the school year in which I gathered data (2001–2002). These interviews afforded me the opportunity to understand whether the patterns I observed in focal teachers’ classrooms appeared in other school contexts.

DATA COLLECTION

Classroom observations, interviews and student work formed the core of my data for this research:

1. Classroom Observations: I spent 68 hours in each focal language arts teacher’s classroom spread across four key points during the 2001–2002 school year before and after testing. On the same day, I
observed each teacher teach two tracks, which allowed me to study how the same teacher implemented instruction across tracks without having to control for teacher background. I took extensive field notes of classroom observations, recording what the teacher and students said and the approximate amount of time spent on components of the lessons.

2. Interviews with Focal Teachers: I interviewed focal teachers six times during a 12-month period (2001–2002). Interviews allowed the opportunity to hear teachers’ thoughts on various issues such as testing, curricular goals and their instructional choices.

3. Student Work: I analyzed a collection of student work that the district required language arts teachers to keep for student promotion decisions.

4. Interviews with Teachers at Other Schools: The teachers who volunteered to participate in the interviews engaged in a one-hour conversation with me about their classroom instruction in the context of the state reform.

These multiple sources allowed me to triangulate my data to warrant my assertions. In addition, documents from the school, district and State Department of Public Instruction, communications with district and state officials, as well as attendance at faculty meetings, district professional development days, and a state “closing the achievement gap” conference further helped me situate the embedded contexts in which the focal teachers worked.

Since I did not research classroom instruction both before and after implementation of the state reform, my research does not position me to comment on the effects of high-stakes accountability reform on classroom instruction. However, my research provides detailed portraits of instruction across tracks in the new context of high-stakes statewide accountability programs. In addition, in order to understand the influences of state policy, I rely on teacher testimony about instruction previous to the state reform and only write about connections between testing policy and classroom decisions when teachers make these explicit links.

CONTEXT OF THE STUDY

SCHOOL DEMOGRAPHICS

Sequoia Middle School was a large middle school, serving 1256 students during the 2001–2002 school year. The school’s demographics were 47.18% White, 46.00% Black, 2.54% Multi-racial, 2.22% Latino, 1.91%
Asian, and 0.16% Native American; 27.15% of the students at Sequoia Middle School qualified for free/reduced lunch (Ridge City Public Schools, 2001a).

SCHOOL CLIMATE: SEQUOIA IN THE CONTEXT OF THE STATEWIDE ACCOUNTABILITY PROGRAM

Sequoia Middle School was considered a desirable school in the district. Teachers described Sequoia as a collegial environment, with high morale, strong leadership, experienced teachers and parent support. Yet the state had assigned Sequoia Middle School a *no recognition* school classification for the past two years. At the time I conducted this research, there were four categories of school classifications in the North Carolina ABCs school reform plan, as Table 1 illustrates; *no recognition* schools were schools that failed to reach the growth standard but had more than 50% of students at or above grade level.

<table>
<thead>
<tr>
<th>School classification</th>
<th>Growth standard</th>
<th>Performance standard</th>
<th>Rewards and Punitive Sanctions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exemplary</td>
<td>Exceeded</td>
<td>Does not matter</td>
<td>$1500 teacher bonus; possible principal promotion</td>
</tr>
<tr>
<td>Expected</td>
<td>Met</td>
<td>Does not matter</td>
<td>$750 teacher bonus</td>
</tr>
<tr>
<td>No Recognition</td>
<td>Not Met</td>
<td>&gt;50% of students at grade level</td>
<td>15 lowest performing schools receive help from state assistance teams; administrators notify parents of status; possible principal demotion</td>
</tr>
<tr>
<td>Low Performing</td>
<td>Not Met</td>
<td>&lt;50% of students at grade level</td>
<td></td>
</tr>
</tbody>
</table>

Sequoia Middle School had slightly improved its performance standard, where the percentage of students at grade level had grown from 76.2% during the 1996–1997 school year to 77.6% during the 2000–2001 school year. However, the school, which received the exemplary school classification from 1996–1998, had a downward school classification trajectory because it did not meet state expectations for growth. The school received a met expectations school classification during the 1998–1999 school year, and the no recognition status from 1999–2001. Though the achievement gap does not factor into the school classification, a 34.3-
point percentage gap remained between Black and White students performing at grade level at Sequoia Middle School in the year preceding the 2001–2002 school year when I conducted my research (North Carolina Department of Public Instruction, 2001). Since 2002, the gap has decreased, but a 24.5% to 29.5% gap still persisted in 2005 (North Carolina Department of Public Instruction, 2005).  

In sum, while the community viewed Sequoia Middle School as a desirable school, educators faced a challenge to meet or exceed the state’s improvement goals during the 2001–2002 school year. This setting provides the backdrop in which Ms. Jones and Ms. Royster taught in their classrooms.

ENROLLMENT PATTERNS ACROSS TRACKS

Looking into the day-to-day classroom contexts of teaching, Black students are disproportionately represented in the “regular” classes and White students are disproportionately represented in the “academically gifted” classes in comparison to the school’s demographics, as Figure 1 displays, across Ms. Jones’ and Ms. Royster’s language arts classes. In addition, economically disadvantaged students are disproportionately represented in “regular” classes in comparison to whole-school demographics, as Figure 2 illuminates.

Figure 1. Racial demographic of classes by teacher (Ms. Jones and Ms. Royster) and track level (“academically gifted” and “regular”).

![Class by Teacher and Level](image_url)
These patterns in the classrooms of the focal teachers at Sequoia Middle School are a microcosm of wider patterns in the state and the country. Mickelson (1999) reports how “tracks are highly segregated with African American students disproportionately in lower tracks and White students disproportionately in higher tracks” in Charlotte-Mecklenburg, another North Carolina school district (p. 8). Lucas (1999) documents an enrollment disadvantage for high school students from low socioeconomic backgrounds in higher academic tracks drawing from a national data set. Oakes (1985) also reports that students of color and students from low socioeconomic backgrounds are disproportionately represented in the lower tracks while children from higher socioeconomic backgrounds are over represented in the higher tracks.

At Sequoia Middle School these student enrollment trends are a product of official and unofficial placement practices. In the official process that the Executive Director for Grants Administration in the district described, students with high IQ scores, as measured by the Test of Cognitive Skills, move on to the next screening level, where educators examine other performance indicators, including achievement test scores. In practice, however, I heard and observed teachers place students into “academically gifted” classes based on their own perceptions of student attitude and effort, their expectations for student achievement based on appearance and cultural background, and parental interventions. These placement practices reflect patterns documented in the research literature, which demonstrate that placement in an ability group is not based on objective measures (Cicourel & Kitsuse, 1963; Lucas, 1999; Mickelson, 1999; Rist, 1973; Rosenbaum, 1976; Schafer, Olea, & Polk, 1972; Turner, 1960). Because track placement is susceptible to parental interventions and subjective decisions by teachers, I have
decided to place the labels “academically gifted” and “regular” in quotation marks in this article to indicate that classroom membership is not based on an objective identification process in practice.

FINDINGS

“I’m sorry you’ll have to see this twice,” Ms. Jones warned me before the start of the second class period I observed for the day (9/10/01 SJ field notes, p. 19). As I observed an “academically gifted” class followed by a “regular” class taught by the same teacher, I was initially struck by the similarity in the units and major classroom activities across tracks. Previous tracking literature had, after all, documented widely different curricula and instruction across tracks in terms of content, skills, intellectual processes, instructional time and materials (Oakes 1985; Page 1989, 1991). When I made a calendar of all of the major instructional units and classroom activities across tracks in my two focal teachers’ classrooms, the curricula looked remarkably similar: Students across tracks read the same novels or short stories or practiced similar skills such as diagramming sentences. However, despite the relative uniformity of the major classroom activities across tracks, I did see differences across tracks when I dug into the nuances of instruction.

In the new school reform context of a high-stakes, statewide accountability program, five key differences in instruction across tracks favored students in the “academically gifted” classes. Students in the “academically gifted” classes had less explicit test preparation, which meant more time to engage in other curricula that teachers deemed important. Students in the “academically gifted” classes additionally received more opportunities to practice a wider range of reading and writing skills, engaged in more challenging instruction and assignments and received more written and immediate feedback on essay assignments than their peers in the “regular” classes. While these patterns became salient in my observations of focal teachers’ classrooms across tracks, these differences in instruction were also evident at other schools, as described by the language arts teachers I interviewed.7

LESS EXPLICIT TEST PREPARATION FOR “ACADEMICALLY GIFTED” TRACK STUDENTS

Students in the “academically gifted” classes received less explicit test preparation compared to their peers in the “regular” classes. As Table 2 illustrates, almost half of the instruction I observed in Ms. Royster’s “regular” class (46.15%) had a high level of correspondence with what the
standardized tests asked students to do by demand (e.g., identify main idea) and format (multiple-choice questions). For example, I observed students in Ms. Royster’s “regular” class work independently on a section of a practice end-of-grade test. In contrast, only 15.38% of instruction I observed for her “academically gifted” class had a high level of correspondence to the standardized tests. Most instruction (61.54%) for students in the “academically gifted” class exhibited a medium level of correspondence with testing demands and format, where students practiced the demands of the test in a format that did not correspond to the test. For example, students in Ms. Royster’s “academically gifted” class read aloud a chapter of the novel, *Waiting for the Rain*, in pairs, and asked each other questions about what they read. Ms. Royster confirmed that students in her seventh period, “regular” class, whom I observed, had more explicit test preparation than students in the “academically gifted” class. She noted, “The only thing I might do differently with seventh (period) is give them more practice “testlets”…” (MJR, Interview 5, p. 8).

**Table 2. Percent of Observed Classroom Instruction by Level of Correspondence with Testing Demands and Format**

<table>
<thead>
<tr>
<th>Level of Correspondence</th>
<th>Regular</th>
<th>Academically Gifted</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>46.15%</td>
<td>15.38%</td>
</tr>
<tr>
<td>Medium</td>
<td>23.08%</td>
<td>61.54%</td>
</tr>
<tr>
<td>Low</td>
<td>30.77%</td>
<td>23.08%</td>
</tr>
<tr>
<td>None</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

1 A low level of correspondence indicated that classroom instruction was tangentially related to the demands of the test. For example, learning to diagram sentences, a demand that neither the writing or multiple-choice tests required, might have helped students with their writing as they learned more about sentence formation.

Interviews with teachers at other schools supported this pattern of findings at Sequoia Middle School. Teachers described pronounced differences in the amount of explicit test preparation across tracks. For students in the “academically gifted” classes, Ms. Mover, a teacher at a school with a high and stable school classification history, “went over one or two examples from (end-of-grade practice books). There are basically six types of questions that they can be asked and we went over those and the differences between them” (JM, River School, p. 13). In contrast, with students in her “regular” classes, Ms. Mover recounted how she “literally went over every example of every type of question in this book which provided at least six or seven reading passages for each type of question” (JM, River School, p. 15). Students in Ms. Mover’s “regular” classes worked from these practice books “three or four days a week” from mid-February when she first got the books until the end-of-grade test in May, while students in her “academically gifted” classes worked from these books.
“maybe an hour or two a week at the most, not very much” (JM, River, p. 19). Students in the “regular” track classes thus did more test preparation than their peers in the “academically gifted” classes, curricula which displaced other priorities teachers valued.

MORE TIME FOR TEACHER PRIORITIES FOR “ACADEMICALLY GIFTED” STUDENTS

For students in the “academically gifted” classes, less explicit test preparation freed time to engage in curricula that teachers across five school contexts deemed important (Watanabe, 2007, a). Specifically, students in the “academically gifted” classes had more opportunities to practice various writing genres. Additionally, they completed more independent assignments where they worked on their communication and collaboration skills and where they developed an appreciation for literature, compared to their peers in the “regular” classes.

Writing Different Genres for Authentic Audiences

Students in “academically gifted” classes engaged in different writing styles, beyond the five paragraph essay the writing test demanded, compared to students in the “regular” classes. Even though Ms. Jones assigned the same novel, The Master Puppeteer by Katherine Paterson, to all of her students across tracks, students in the “academically gifted” (AG) class read the last several chapters of the novel independently at home during the Thanksgiving holiday, allowing them to do other assignments upon their return to class. For example, students in Ms. Jones’ “academically gifted” class wrote haikus and stories that they read to children at a nearby elementary school while students in the “regular” class continued to finish The Master Puppeteer with Ms. Jones in class.

One can imagine how an authentic audience could boost students’ motivation to write well in comparison to writing for a test scorer. In addition, students in the “academically gifted” class practiced skills for the writing test in a “fun” context (SJ, Interview 3, p. 4). For example, they practiced their organization, coherence and written language conventions skills, which are part of the evaluation process for the writing test, in order to communicate a story to children. Instructional opportunities to enjoy writing are important considering teachers’ fears that their students view writing as “boring” with all of the writing test practice students complete (JS, Cornwallis, p. 11).
More Independent Projects – Instructional Opportunities to Appreciate Literature and Work on Communication and Collaboration Skills

Independent projects, which teachers appeared more likely to assign to their students in the “academically gifted” classes than to students in their “regular” classes, provided opportunities to read and appreciate literature and work on communication and collaboration skills. For example, while Ms. Mover guided students in her “regular” class through explicit end-of-grade test preparation, she allowed students in her “academically gifted” class to read another novel during that time...(and do) a novel project, where basically...all my students have a lot of choices of what they (can) do—some of them build models, some of them act out some of the scenes. During another part of that time period we had a couple of debates and...my kids did public speaking on a specific topic that they researched. And then we (also) did a talk show format where they played different roles... (JM, River School, 15)

Students in Ms. Mover’s “academically gifted” classes thereby practiced their communication skills through public speaking and developed their appreciation for literature by building models using texts as resources. These projects further helped build students’ collaboration skills as students worked together in groups to act out scenes in literature.

While I did not observe or often hear teachers describe sociocultural approaches to literacy, when it occurred, students in the “academically gifted” track had more opportunities than their peers in the “regular” classes to interpret literature with each other without the teacher asking the discussion questions. For example, Ms. Blair, a teacher at a school with an initially low school classification, that since demonstrated an upward trajectory, spoke of how she did more Paideia seminars with her advanced class, “where...we’re having a literary discussion about (a novel) and it’s open—there’s no raising of hands in Paideia and they’re free to draw from their personal experiences, to infer things that are happening in the story” (SB, Rose Valley, p. 16). In sum, students in the “academically gifted” classes had more opportunities to work on their communication skills and to share their appreciation for literature as they built on each other’s understanding of literature through projects and Paideia seminars.

These key differences in instruction between the “regular” and “academically gifted” classes reflect similar dynamics reported in earlier
tracking literature. Oakes (1985) found the “teaching of reading skills, generally by means of workbooks, kits and reading texts,” characteristic of low track classrooms (p. 76). In contrast, she reported how “teachers of high-track classes were more likely to emphasize such behaviors as critical thinking, independent work, active participation, self-direction and creativity” than teachers who taught low-track classes (p. 85). On one level, the descriptions of classroom instruction illuminated in this research are consistent with Oakes’ work; students in the “regular” class did more isolated reading skill exercises compared to their counterparts in the “academically gifted” classes in the new context of high-stakes statewide accountability programs. The consistency of classroom dynamics in the new policy context is an interesting finding in and of itself; despite the fact that North Carolina policymakers herald the high-stakes accountability program as an important component in their efforts to close the achievement gap by expecting high achievement from all students, the state accountability policy did not promote challenging and equitable instructional opportunities across tracks.

Moreover, in this research, teacher testimony about instruction before and after high-stakes testing and about what they would do differently without high-stakes testing suggested that high-stakes testing has changed the ways in which teachers approach instruction across tracks. Namely, high-stakes testing has influenced teachers to do more isolated reading skill instruction for all students, but particularly for students in the “regular” classes, through explicit test preparation, whereby students practice the demands and format of the multiple-choice standardized test through workbook exercises. In a context without high-stakes testing, teachers noted that they devoted more time to a wider range of reading and writing instruction with more independent and authentic assignments for students across tracks. For example, Ms. Greenleaf, a teacher at a school with an initially low school classification that since demonstrated an upward trajectory, used to assign independent poetry projects whereby students individually picked a poet, researched his/her life, identified a favorite poem to paraphrase, described its rhyme scheme, illustrated the poem and wrote a parody of the poem. In a high-stakes testing climate, however, Ms. Greenleaf noted, she “couldn’t do it all” and she abandoned this project for her students in the “regular” classes to pursue explicit test preparation (SG, Rose Valley, p. 5).

Thus, the fact that there continued to be a disparity in instructional opportunities across tracks is consistent with previous literature; however, the nature of this gap has changed. The new context of high-stakes statewide accountability programs appears to have shifted the instructional focus to emphasize isolated reading skill instruction for all
students across tracks, but particularly for students in the “regular” track, at the expense of other curricula that teachers deemed important, such as working independently and collaboratively on creative literature-based projects, appreciating literature, and enjoying writing a variety of genres. The state accountability program did not interrupt the pattern of inequitable instructional opportunities across tracks, documented in previous tracking literature. Therefore, while high-stakes statewide accountability programs may be presumed to increase academic rigor for all students theoretically (Skrla et al., 2000), the accountability program did not encourage teachers to question inequitable instructional practices across tracks or tracking as an organizational practice in the first place.

State policymakers did not appear to consider the possible policy collision between its accountability policy, part of its efforts to close the achievement gap by expecting high achievement from all students, and its tracking policies (SB1207, Article 9), which require local school boards to create distinct programs for its academically and/or intellectually gifted students. The state’s tracking policy sets up the expectation that students in the “academically gifted” track should engage with more challenging content and enriching instructional approaches than their counterparts in the “regular” track, and reinforces teachers’ decisions to offer students across tracks inequitable instructional opportunities. The bifurcated curriculum, which the state’s tracking policy promotes, serves to widen the gap rather than close it, as lower-track students (predominantly African American students) spend more time on test practice and test preparation than students in the “academically gifted” classes, who are more likely to be White and more economically privileged, and who engage in more substantive content and skills.

While high-stakes testing influenced differential treatment across tracks discussed thus far (less explicit test preparation and more time for teacher priorities for “academically gifted” students), high-stakes testing did not appear to influence teachers’ decisions to teach differently across tracks for the three other differences that follow.

MORE READING AND WRITING PRACTICE: WORK AT HOME AND IN CLASS

Students in the “academically gifted” classes completed more reading and writing assignments compared to their peers in the “regular” classes. One reason for this difference in quantity of assignments was that teachers assigned more reading and writing work to students in the “academically gifted” classes to complete at home. For example, by the end of the school year, students in Ms. Royster’s “academically gifted” class read two
more novels than students in her “regular” classes because Ms. Royster assigned these novels to students to read outside of school. Ms. Jones also asked students in her “academically gifted” class to finish essay assignments at home after she started them on the assignment in class and “monitor(ed) (to) see that they (were) on the right track.” In contrast, with students in her “regular” classes who “are not as readily available without prompting and structure, or may not have equal help at home, I may have them write the introductory paragraph (in class),” Ms. Jones explained (SJ, Interview 4, p. 6). This finding is consistent with the patterns that previous researchers have described about instruction across tracks in a context without high-stakes accountability programs. Page (1989) reported that teachers did not assign students in low-track classrooms as many assignments outside of class compared to students in high-track classrooms.

While policymakers and educators may note that it seems reasonable that more “advanced” students read and write more than “less advanced” students, the extra novels and writing assignments provide more opportunities for students in the “academically gifted” classes to practice their reading and writing skills for the state tests compared to students in “regular” classes. In addition, this “story” is not just about the differential quantity of reading and writing opportunities across tracks. It is also about HOW teachers teach the same novel to different groups of students, affording different opportunities for learning.

More Actual Reading, as opposed to Storytelling, in “Academically Gifted” Track Classrooms

Due to the differences in the way in which Ms. Jones taught the same novel across tracks, students across classes also received different opportunities to practice their reading skills. In comparison to students in the “regular” class, who were more likely to listen to Ms. Jones tell the story, students in the “academically gifted” class engaged with the text as they read and discussed the novel in class. The following descriptions of classroom instruction clarify this difference in Ms. Jones’ approach to teaching the same chapter across tracks.

One day, I observe Ms. Jones read Chapter Seven of the Master Puppeteer to students in her “academically gifted” class at a fast pace, shouting out the Japanese word, “Ara,” expressing surprise, as if she herself were the book’s protagonist, Jiro. Jiro is a poor, young, Japanese boy, who leaves his family to learn the artistry
of bunraku, or Japanese puppetry, at the Hanaza, a famous puppetry theater, during a time of social upheaval and poverty in 18th century Osaka. Ms. Jones suddenly stops reading. “Houston, we’ve got a problem,” she notes. She summarizes events in the last chapter that lead to the conflict. Jiro leaves the Hanaza on New Year’s Day to deliver food to his starving mother. On his return to the Hanaza, an outlaw samurai (ronin) threatens Jiro with his sword, but a man fells the ronin in one swoop, saving Jiro. Jiro is surprised to see that the man who has saved his life is his stern master teacher, Yoshida, whom Jiro thinks is sick in bed. Yoshida threatens Jiro not to tell a soul that he has seen him, and Jiro hurriedly returns to the Hanaza.

In the process, Jiro has “forgotten to make the pilgrimage to the shrine, to the church. If you want fortune to smile upon you, you go to the shrine. If you forget (the Gods) they might forget you,” Ms. Jones explains the Japanese tradition of visiting a shrine on New Year’s Day. “Why does he not go back?” Ms. Jones asks. Patrice, an African American student, says that Jiro does not want to see Yoshida. Ms. Jones confirms and continues to read on.

When Jiro returns to the Hanaza, he sees his fellow apprentices drunk, celebrating the New Year. When Ms. Jones reads the parts of the apprentices, she stutters as is written in the text, “‘You-you missed all the fun.’ What have they been doing?” Ms. Jones asks the class. Wilson, a White student, calls out that the apprentices have been drinking sake. “‘S-sspeak for yourself, piglet,’ ” Ms. Jones amuses the students as she slurs her speech, acting as if she herself were drunk. “Why was it obvious that Yoshida is not there?” she asks. Kelly, a White student, says the apprentices are having a party and Ms. Jones says, yes, “the boys are loose. They’re just whoop-de-doing.” After Ms. Jones continues to read and discuss the text a bit more, she asks Christopher, a White student, to read. Christopher reads a section about the assignment of parts to the apprentices for the new play. Ms. Jones then asks students to think about why the apprentices are unhappy with their parts (11/12/01 SJ field notes, p. 5–6).

In this description of classroom instruction, Ms. Jones dramatically reads the text aloud, often acting out the characters’ parts. She asks students reading comprehension questions and provides her interpretation.
of the events. She then calls on Christopher to read. This “academically gifted” class is 85.3% White, and 2.9% of students qualify for free and/or reduced lunch.

In contrast, Ms. Jones teaches the same chapter to her “regular” class quite differently. This class is 73.7% Black, and 26.3% of students qualify for free and/or reduced lunch. The times and specific names of students in the following description of classroom instruction are significant in understanding how much time particular students have to read the text independently.

At 12:46 pm, Ms. Jones tells her students in the “regular” class to take a short quiz on Chapter Six and when they finish, to read Chapter Seven individually. At 1 pm, Jan, one of the few White students in class, has finished the quiz and reads, her shoulders hunched over the book. “I’ll give you a worksheet with questions as a reading guide for Chapter Seven,” Ms. Jones says. She adds, “It will help you know what’s important.” At 1:04, half of the students are working on the quiz and the other half are reading.

At 1:12, it is still quiet. With the exception of Kiona, a Black student, everyone is reading Chapter Seven. At 1:13, Kiona is finished. Ms. Jones collects her quiz.

At 1:14, Ms. Jones interrupts the silence by beginning to go over the seventh chapter. Ms. Jones pretends to be Jiro, gasping for breath. “What has he just experienced?” she asks. Students laugh at her acting. Sherice, a Black student, summarizes that he has just run back to the Hanaza after his encounter with the ronin. Ms. Jones shouts “BAMN!,” startling a few students out of their seats as she dramatizes when Yoshida stops the ronin from attacking Jiro. Jiro “was like a rabbit jumping out of his hole. (But) he has forgotten something,” Ms. Jones pauses dramatically. Donté, a Black student, responds, “A pilgrimage?” Ms. Jones confirms, then pretends to be Jiro, “I’ll go now, no I won’t go.” “Why does he change his mind?” Ms. Jones asks. After further discussion of Jiro’s predicament and the description of the drunk apprentices at the Hanaza, the bell rings, and further teaching of Chapter Seven is tabled until the next day (11/13/01 SJ field notes, p. 4–7).

As students file into class and drop their heavy backpacks on the ground before taking their seats the next day, Ms. Jones asks
Ms. Jones reads the first sentence of the chapter and says, “Raise your hand if you can tell me where Jiro has just come back from.” Jan says that he was just in the alley with Yoshida. Ms. Jones agrees, then goes on to summarize, “what the whole next paragraph says is this. . . ” She explains that Jiro realizes that he has forgotten to go to the New Year’s shrine. “Doggone it, why did Jiro not stop to pray when he went to his mother’s house?” Ms. Jones asks. Jason, a Black student, responds that Jiro was carrying the soup. Ms. Jones demonstrates how Jiro held the container of soup underneath his clothes, “and he walks bit by bit so he will not draw attention to himself.” Ms. Jones then discusses Jiro’s trip back to the Hanaza. “Why does Jiro not stop at the shrine on his way back?” she asks. Jan answers again, saying that he doesn’t want to see Yoshida. Ms. Jones confirms, and imitates Yoshida threatening Jiro to go back to the Hanaza in a deep voice, “I’m not just whistling Dixie.”

Ms. Jones directs students to the bottom of p. 77 of the novel, where Jiro gets back to the Hanaza. She then asks, “on p. 78, what does Jiro notice?” Hannah, a White student, responds that the apprentices are drunk, and Ms. Jones pretends she is one of the apprentices as she laughs like Daffy Duck, “Haha——-h.” Ms. Jones summarizes that Yoshida has not been back for three days and the apprentices are having a good time. She reads a few sentences from the text where an apprentice named Wada says, “With a little luck (Yoshida’s) sickness will last through the holidays,” and Kinshi says, “with a little luck it may prove fatal.” “What’s fatal?” Ms. Jones asks. Jan responds, “he may die.” Ms. Jones tells her students that “that is blatant disrespect. Usually, Jiro may not mind this talk but why does Jiro not laugh this time?” Jan points out that Yoshida saved his life. Ms. Jones agrees, then directs students to the bottom of p. 78. She tells students that the apprentices start practicing for a new play.

Students across and within tracks in Ms. Jones’ classes have different opportunities to read the text and answer reading comprehension
questions independently on the days on which I observed. Ms. Jones talks through or summarizes Chapter Seven of the *Master Puppeteer* to students in her “regular” class to get the content across; in contrast, she reads the chapter, paragraph by paragraph with students in her “academically gifted” class. Talking through vs. reading the text presents different opportunities for students to practice reading and interpreting text, demands students must perform on the standardized reading test.

Within the “regular” track class, students who complete quizzes more quickly get more practice with reading; some students do not have any opportunity to read the text at all. For example, on the first day of Chapter Seven instruction, Jan finishes the quiz and begins reading at 12:58. Ms. Jones starts discussing Chapter Seven at 1:14. In contrast, Kiona, who was still working on the quiz at 1:12, does not have ANY time to read the chapter independently before Ms. Jones begins to retell the major events in Chapter Seven. It is also interesting to note that Jan, a student who has had more time than her peers to read Chapter Seven independently, answers four of Ms. Jones’ six oral comprehension questions in the portrait of classroom instruction on the second day. She is one of the few students who has had the opportunity to read the text, thus enabling her to answer the teacher’s questions.

Ms. Jones’ instructional choices may appear puzzling to readers; however, Ms. Jones’ perspective helps shed light on why she chooses to teach the novel in this way. Ms. Jones explains that the *Master Puppeteer* is above the reading level of many students in her “regular” classes. However, she teaches the novel because of its tie-in with the social studies curriculum and students’ appreciation of the novel. She comments,

(The social studies teacher)…has lived in Japan…and she goes into the puppetry. She has a movie on (a)Yoshida (who makes puppets), so as I’m reading about Yoshida (students) come to me and say oh, we (saw)...Yoshida...And there’s so much she has to offer that if she offered it to everybody and one group had only read the *Master Puppeteer*, then the other three groups would not appreciate everything she had to offer as much

(SJ, Interview 3, p. 9-10)

Ms. Jones also finds that students in her “regular” classes appreciate reading the novel. Ms. Jones describes the reaction of her students:

As the story (goes) on, the tension mount(s) and by the end, (students in my ‘regular’ classes are) saying, ‘oh gosh, we’re glad we read this book. This was good, Ms. J.’ And it (makes) me think
(that) the pulling of . . . teeth and the blood, sweat and tears (is) maybe worth it. Plus it (makes) them feel special that they (are) reading what the ‘AGs’ (are) reading (SJ, Interview 4, p. 15).

While Ms. Jones feels that it is important for her to teach the novel for these equality-based reasons, she faces the difficult challenge of teaching a novel written at a higher reading level than the reading levels of some of her students.

Ms. Jones retells the story more in the “regular” class than in the “academically gifted” class primarily because she feels the reading level is just too difficult to handle for many of the students in her “regular” class. She notes, “Well, the thing about it is, (the “academically gifted” class) can read and understand some of these things. . . it’s the caliber of child I have” that facilitates her teaching the Master Puppeteer (SJ, Interview 3, p. 7). In contrast, she has the “regular” class:

…read along as I read. Now when the paragraphs (are) short I…call on them but that way I (can) talk about Japan at the time of the samurai warriors. I (can) discuss each paragraph before we (go) on to the next so that I…know that I (have) them with me (SJ, Interview 4, p. 15).

Ms. Jones feels it is necessary to read the longer passages aloud and scaffold reading comprehension in her “regular” classes.

While Ms. Jones exposes students in her “regular” class to curricula integrated with social studies content, these students do not practice how to approach difficult text, a skill that could help them read challenging end-of-grade test passages. When I asked Ms. Jones “where…these students get their reading practice…if these novels and these textbooks are all at such a different level than their reading level,” Ms. Jones responded, “That’s a good question, because I don’t know that the math and the science and the social studies (teachers) really use their textbooks, either…A lot of money wasted” (SJ, Interview 4, p. 15). Thus, even though Ms. Jones decides to teach the same novel to both classes to engage students in similar content, the way in which she teaches the novel affords her students very different opportunities to read and practice reading comprehension questions.

Finally, Ms. Jones’ rationale for the instructional difference across tracks makes it clear that the high-stakes accountability context is not influencing her to teach across tracks differently. Instead, she reports that her assessment of students’ reading and writing levels, their motivation and availability of help to complete work at home influence her to
assign more independent reading and writing work and more actual in-class reading and analysis, as opposed to storytelling, for students in the higher track. The new accountability context appears not to have challenged the prior beliefs of well-intentioned teachers like Ms. Jones to consider the implications of their instructional choices for the development of students’ reading and writing skills.

ONE STEP AHEAD: MORE CHALLENGING INSTRUCTION AND ASSIGNMENTS

Teachers taught more challenging material to students in the “academically gifted” classes compared to students in the “regular” classes. This decision appeared to stem not from the high-stakes accountability program but from teachers’ belief in a linear progression of skills. Teachers believed that students in the “academically gifted” classes had mastered basic skills and were capable of moving on to higher-order skills. Therefore, they assigned novels at a higher reading level, higher-order discussion questions and more challenging essay topics linked to current events to students in their “academically gifted” classes.

*Novels at a higher reading level for students in the “academically gifted” track*

Language arts teachers sometimes assigned students in the “academically gifted” classes novels at a higher reading level compared to students in “regular” classes, while making efforts to ensure that the different novels took place in the same country students studied in their social studies classes. For example, both Ms. Royster and Ms. Jones assigned students in their “regular” classes the novel, *Journey to Jo’Burg*, while assigning students in their “academically gifted” classes the novel, *Waiting for the Rain*. Ms. Jones noted that both novels discuss the experiences of children living under apartheid in South Africa in the 1970s, thereby providing the opportunity to talk about similar issues across the different levels of her classes. However, Learning Links, Inc., a publisher of both books, notes that *Waiting for the Rain* is written at the fifth through eighth grade level while *Journey to Jo’Burg* is written at the third through fifth grade level (Books in Print, 2002). This finding is consistent with previous research in a context without high-stakes testing that documented how students in low track classrooms read work that is “written at a low level of difficulty” (Oakes, 1985, p. 76).

Teachers face a dilemma in their decision to assign novels at different reading levels to students. On one hand, teachers want to afford students the opportunity to practice reading at a level that is not overwhelmingly
difficult for them. On the other hand, the selection of reading material at different levels over time can solidify or even widen the gap in performance across groups of students if teachers only expose their students to material at students’ ability levels. In addition, the assignment of novels at different reading levels provides divergent opportunities for students to practice skills for the high-stakes state reading test, depending on the correspondence between the reading levels of the assigned novels and test passages. Test developers select the end-of-grade test passages to provide a means to evaluate the “grade-level” performance of students, in this case, that of seventh graders. If teachers expose students to strategies to tackle difficult text in a novel that is similar in reading level to the end-of-grade test reading passages, students can presumably apply these strategies to difficult reading passages on the end-of-grade (EOG) test. Conversely, if teachers do not expose students to text that is similar in reading level to the EOG reading passages, as would be the case when students in the “regular” classes read *Journey to Jo’Burg*, students may not be equipped to approach difficult text material. Teachers’ choices for novels ultimately provide different opportunities to students across tracks to approach passages on the EOG test.

*More higher-level questions for students in the “academically gifted” track*

Teachers also created more opportunities for students in their “academically gifted” classes to delve into higher-level questions about the short stories and novels they read, through discussion question assignments and in-class discussions, compared to students in their “regular” classes. Thus, “academically gifted” students not only experienced a rich forum for learning but also gained more practice for the state standardized tests that demanded that students be able to “draw conclusions,” “make inferences” and “read critically for literal and interpretive meaning” (Ridge City Public Schools, 2001b, p. 8).

1. More Discussion Question Assignments

Both Ms. Royster and Ms. Jones assigned discussion question assignments, which Ms. Royster described as

writing assignments that gear (students) towards the higher level thinking, the critical thinking, the inference, the drawing conclusions. It’s not as much recalling information . . . or sequencing the information. . . You need to tell me what kind of conclusion can you draw based on what’s in the text, what the author
Discussion question assignments thus provided an opportunity for students to deeply engage with the text.

For example, the day after Ms. Jones assigned written discussion questions to students in her “academically gifted” class on the novel, *Waiting for the Rain*, Ms. Jones calls students to what she playfully terms the “hot seat,” a seat at the front of the classroom, where students individually read their written responses. When the class discusses the question, “Joseph (a Black South African teenager) refers to the White liberals. What is his personal opinion of these people? Support the answer given with specific reasoning,” students build on each others’ understanding of the material. The discussion reaches a critical moment with Lucy’s contribution. Lucy, a White student, says that initially she thought that Joseph did not hate the White liberals but when she reexamined the book it made her “rethink that.” Lucy continues that when Joseph’s cousin, Tengo, expresses appreciation that Dr. Miller, a White liberal, paid for his book fees, Joseph comments that the book fees are “much less than the fees at the golf club.” Lucy says that the text notes that Joseph said that comment “shortly.” Ms. Jones affirms Lucy’s interpretation, commenting that readers can hear the bitterness in that comment (4/18/02 SJ field notes). Through these written discussion questions, students in Ms. Jones’ “academically gifted” class practice supporting an argument with details, a skill which the point-of-view prompt on the writing test demands. Students listen to each other’s writing, build on each other’s responses, and learn from classmates, like Lucy, who demonstrate how to make arguments based on information presented in the text.

While students in Ms. Jones’ “regular” class orally discussed similar issues in their novel about South Africa, *Journey to Jo’burg*, they did not write out their answers as much as students in the “academically gifted” class. Ms. Jones explained that “one of the reasons is, that the ‘regular’ classes can’t get as much done in a period and you want to cover the whole book and go on to the next book and keep them up with the others, so that they can get at least the same (number of) novels” (SJ, Interview 6, p. 8). Ms. Jones further did not ask students to write out answers at home, because she has noticed that “not a lot is going to get done at home with the ‘regular’ end of the spectrum” (SJ, Interview 6, p. 9).

Ms. Royster similarly suggested that students in her “regular” classes completed discussion questions that “geared (students) toward the higher level thinking” less than students in the “academically gifted” class.
Students in both focal teachers’ “regular” classes thus had fewer opportunities to experience the type of classroom instruction that students like Lucy benefited from, and to practice the higher-order thinking skills that the end-of-grade tests asked them to demonstrate.

2. In-Class Discussions

Moreover, while Ms. Royster often taught the same short stories and novels across tracks, she posed different types of questions to students during in-class discussions in the “regular” and “academically gifted” classes that prepare students differently for the demands of the state standardized tests. Ms. Royster broke down the steps for students in her “regular” class to think about broader discussion questions that she posed directly to students in her “academically gifted” class.

The field note excerpts in Table 3 demonstrate how Ms. Royster approaches discussion differently across tracks about the cultural differences between Amy Tan’s Chinese family and an American minister’s family, as the two families sit together for a Chinese meal. I italicize the questions Ms. Royster poses to highlight the difference in the questions she asks. In these excerpts on cultural differences at the dinner table, Ms. Royster asks students in her “academically gifted” class to contrast Chinese and American table manners and to infer from the story what is acceptable in Chinese culture. She starts with broader questions in comparison to what she asks students in her “regular” class. With students in her “regular” class, Ms. Royster poses a detail question, “What are they doing, guys?” and then follows up by explicitly asking students to contrast one aspect of Chinese culture with American culture, “Is that something we do?” In addition, instead of asking students to infer from the text what is acceptable in Chinese culture as she did with students in her “academically gifted” class, Ms. Royster tells students in the “regular” class that in Chinese culture you do not have to wait for the dishes to be passed. Ms. Royster breaks down the issue about cultural differences at the dinner table for students in her “regular” class.

Ms. Royster is aware that she asks different types of questions to her students across tracks. She explains,

I don’t know if you noticed, (but)…I phrase my questions differently (across classes)…(I) probably do more repetition with (my regular students) in that I may phrase a question the way I would phrase it for my AG student and then before somebody answers it I turn around and give a simpler version of the question and
then I may even turn around and give an even third version of the question, so they’re hearing different things.
(MJR, Int 2, p. 4–5)

While Ms. Royster is aware that she phrases questions differently across tracks, my analysis of the larger data set, including the data excerpt presented in Table 3, suggests that Ms. Royster does not initially “phrase (the) question the way (she) would phrase it for (her) AG student(s).” Instead, she is more likely to begin a question sequence that moves students in her “regular” class step by step to consider the broader question she poses to students in her “academically gifted” class.

Table 3. Excerpts from instruction that display the differences in the ways in which Ms. Royster approaches discussion topics across tracks.

<table>
<thead>
<tr>
<th>Tracks</th>
<th>“Academically gifted”</th>
<th>“Regular”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultural Differences at Dinner Table</td>
<td>Ms. Royster asks, “What are some of the cultural differences taking place at the dinner table?” Heather, a White student, says that the Chinese family is using chopsticks, and that family members are eating the food and serving themselves with the same utensils.</td>
<td>Ms. Royster reads the sentence about how Amy’s Chinese relatives use the same utensils in their mouths and in the shared dishes. “What are they doing, guys?” she asks. “Spreading bacteria,” Henry, a Black student, responds. “Is that something we do in American culture?” Ms. Royster asks. “No ——,” several students say simultaneously. “How many of your mamas would hit you if you did that?” Ms. Royster asks. Several hands go up in the air.</td>
</tr>
<tr>
<td></td>
<td>Ms. Royster points out that the minister’s family waits for the platters patiently. “Based on what it says in the text, what does this say about Chinese culture?” Brian, a White student, responds that the Chinese don’t pass platters around. Ms. Royster affirms this answer and notes that in the background section of the text on p. 76, it explains how the Chinese eat “family style.”</td>
<td>Ms. Royster then directs the class’ attention to how the minister’s family patiently wait for the platters to be passed to them. She then asks, “Were platters ever passed to them?” “No, they were hogging food,” “passing hepatitis,” students call out.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ms. Royster says that in Chinese culture, you “go for it” together instead of waiting for dishes to be passed. Jerome, a Black student, responds, “Same thing sometimes for us. It depends.” Adrian, a Black student, adds that that is what you do when you go to one of those buffet dinners or a cookout. “Well that’s different, of course,” says Ms. Royster.</td>
</tr>
</tbody>
</table>
This analysis illustrates how Ms. Royster provides students in her “academically gifted” class with more challenging questions compared to students in her “regular” classes. Specifically, students in Ms. Royster’s “academically gifted” class get more practice with comparing and contrasting, “read(ing) critically for literal and interpretive meaning,” as well as “mak(ing) inferences” based on information presented in the text, demands that students must demonstrate on the end-of-grade reading test (Ridge City Public Schools, 2001b, p. 8).

3. Essay Topics.

Students in Ms. Royster’s “academically gifted” class also received more opportunities to consider challenging topics linked to current events in their writing compared to students in the “regular” class. For example, Ms. Royster assigned additional essay topics to students in her “academically gifted” class that went beyond the “basic” essays that students in both “regular” and “academically gifted” classes completed. For example, students in the “academically gifted” class wrote about cloning, animal testing, religion in schools, and whether or not the World Trade Center should be rebuilt.

Ms. Royster explained that these “prompts required a little bit more of the higher level thinking because they’re AG students. . . and it was obvious that they could write. They had the basics down for writing” (MJR, Interview 4, p. 4). Instead, Ms. Royster said that she “may do more graphic organizers for some of the other classes than I do for the AG class” (MJR, Interview 4, p. 6). In this quote, Ms. Royster indicates her belief in a sequential theory of learning. According to Ms. Royster, students should have the “basics down” before considering more challenging topics in their writing that required “higher level thinking.”

Ms. Royster further explained that another reason that students in the “academically gifted” class completed “a lot more prompts” compared to students in the “regular” classes was because students in the “academically gifted” class sometimes wrote “in groups…they got together, discussed the essay, did their graphic organizer chart and basically wrote the essay—it was maybe four or five of them together” (MJR, Interview 4, p. 4). Therefore, students in the “academically gifted” class not only received more opportunities to consider challenging topics related to current events in their writing, but also engaged in more socially situated learning activities, where they shared interpretations and built upon each other’s knowledge. Compounding this analysis is that White students are disproportionately overrepresented (72.4%), and Black students (24.1%) as well as students from low socioeconomic backgrounds (6.9%)
are underrepresented in Ms. Royster’s “academically gifted” class, in a school where 47.18% of students are White, 46.00% of students are Black, and 27.15% of students qualify for free and/or reduced lunch. Thus, overall, Black and White students, as well as students from middle to high and low socioeconomic backgrounds experienced very different instruction with respect to the types of discussion questions they encountered in their classrooms.

It is again clear through these teachers’ explanations of differences in instructional practice that teachers’ sequential theories of learning, and not high-stakes testing, are driving teachers to provide more challenging instruction and assignments to students in the higher track. The patterns of instruction highlighted in this research are consistent with previous literature on differential learning opportunities in tracked classes in a context without high-stakes testing. That is, Oakes (1985) documents how instruction in the higher tracks “demanded critical thinking, problem solving, drawing conclusions, making generalizations or evaluating or synthesizing knowledge” as illustrated by the “hot seat” discussion in this research (p. 76). In direct contrast, instruction for students in low-track classrooms emphasized simple comprehension. Page (1989) described how teachers of low-track classrooms often engaged in a “triptartite recitation structure of teacher-question/student-response/teacher-reaction” for “closed” as opposed to open-ended questions, as was evidenced in the discussion about “Fish Cheeks” by Amy Tan in Ms. Royster’s “regular” class (p. 60). In sum, high stakes accountability programs have not disrupted teachers’ beliefs in a linear progression of skills, which appear to drive these teachers’ decisions to teach across tracks in contrasting ways (more reading and writing practice and more challenging instruction and assignments for students in “academically gifted” track).

In addition, absent from focal teachers’ discourse about their rationales for their bifurcated curricular decisions was any mention of tracking policies. In this district, tracking is so deeply embedded in the organization of schools that teachers appeared to treat the existence of the two track levels as matter-of-fact; they did not once mention tracking as an influence on their decision to teach students in “academically gifted” classes more challenging content. Even though these teachers did not identify the state’s tracking policy as an influence on their curricular decisions, teachers’ actions are shaped by the structural organization of schools. The state’s tracking policy serves to reinforce not only teachers’ belief in a linear progression of skills, but also supports their enactment of bifurcated curricula. The state’s tracking policy, which sets up the expectation that students in “academically gifted” programs should be engaged in more challenging content, sanctifies different expectations,
instructional approaches and content for students across tracks, even when this differentiated curricula leads to unequal access to knowledge. What is especially pernicious about the state’s tracking policy is that it is so widely accepted as an organizational practice by educators participating in this research that its influence seems invisible. Despite egalitarian claims, the accountability program does not overcome the stratifications of access to knowledge built into the system of tracking, which are so embedded in the school culture that teachers scarcely identify tracking policies as a piece of what influences their curricular decisions.

MORE WRITTEN AND IMMEDIATE FEEDBACK ON ESSAY ASSIGNMENTS FOR STUDENTS IN THE “ACADEMICALLY GIFTED” CLASSES

Some of the nuanced differences in instructional practices across tracks were not evident initially through classroom observations or interviews alone. An in-depth examination of student work uncovered that the two language arts teachers also gave more written and immediate feedback on essays to students in their “academically gifted” classes compared to students in “regular” classes. To evaluate the written feedback that students across tracks received on their individual essays, I analyzed a collection of student work from the entire school year for students in the two periods of classes I observed for each teacher. In one analysis of essay assignments from Ms. Royster’s classes, I randomly sampled five students’ manila folders from her “academically gifted” and “regular” classes, using a table of random numbers. I then calculated the proportion of essays on which Ms. Royster gave written feedback for each manila folder. Students in the “academically gifted” class received four times as much written feedback as students in the “regular” class. Specifically, Ms. Royster wrote comments on 53.2% of the essays in “academically gifted” students’ folders and 13.2% of essays in “regular” students’ folders.

Each writing folder from Ms. Jones’ classes across tracks contained the same three essays, enabling me to compare the number of written comments or corrections across tracks. Ms. Jones made a total of 275 corrections on the essays of five randomly sampled “academically gifted” students; in contrast, she made a total of 58 corrections on the essays of the five randomly sampled “regular” students. Thus, students in Ms. Jones’ “academically gifted” class received five times as much feedback as students in the “regular” class. Moreover, each of the “academically gifted” students received at least 41 corrections. In Ms. Jones’ “regular” class, however, two out of the five students in the sample received no feedback on their essays.
In addition to receiving a higher quantity of feedback, students in the “academically gifted” class received more immediate feedback on their essays compared to students in the “regular” class. For example, as Ms. Jones returned essays to students in her “academically gifted” class one day before a district-wide practice writing test, she noted, “Yours are about the only ones I finished” (9/10/01, SJ field notes, p. 8).

On one hand, the large number of students teachers taught everyday as well as the lack of structural support built into the school day for teachers to grade constrained teachers from providing written feedback to students. The language arts teachers I observed taught 120+ students everyday, making it very difficult to keep up with the sheer volume of essays from all of their students, especially when they assigned several essays in a short amount of time. In addition, though teachers had two “planning periods” each day, “planning period” was truly a misnomer, as I observed teachers attend to other responsibilities such as meeting with other teachers or parents, organizing fieldtrips, and filling out paperwork.

While the large number of students and lack of time to grade during the school day constrained teachers from providing feedback, these factors did not explain why one group of students received so much more feedback compared to another group of students. When I asked teachers to help me understand the difference in feedback I observed, Ms. Jones noted that

(The ‘AG’ students) wrote more, for one thing…And they put meat between the bones. And they are also not as uptight about criticism. With the others, you pick and choose what you want to say because if their whole page is red they’re just going to ball it up and throw it away, and so you’ve got to kind of do it in a very positive way for kids who don’t have that much self-esteem (SJ, Interview 6, p. 10)

While Ms. Jones first cited the quantity of writing, her explanation does not account for the difference in feedback across classes as I compared the number of comments across tracks on the three essays that were similar in length. The average number of words in the essays of students was 278 for students in the “academically gifted” track and 249 words for students in the “regular” track. Students’ readiness to take criticism, while certainly an important consideration in developing the writing skills of students, also does not help explain why two of the students in the “regular” class received no written feedback on their essays.

Ms. Royster suggested that the amount of feedback across tracks may have varied due to the particular components of the essay she
emphasized at the time students wrote the essays. She elaborated,

It may just be where we are in the writing unit. If this is the first or second essay that we’re working on and they’re editing and revising, then I’m looking for maybe four or five basic things...You know, are you expressing your three reasons in that first paragraph?...And it gets more elaborate as the year goes on...it may have been more so what exactly was the piece that we were working on...? Because the feedback became more elaborate as we went on because I expected more from them.

(MJR, Interview 6, p. 15)

In other words, Ms. Royster may have provided comments on a few components of the essay when she first started to teach writing. As the year went on, she may have written more as her expectations grew. Ms. Royster supposed that I had compared “regular” students’ essays from the beginning of the school year and “academically gifted” students’ essays from the end of the year. However, this explanation does not account for the difference in written feedback across tracks, as the essays I analyzed were written by students across tracks at about the same time during the school year.

Teachers thus appeared to be unaware of the extent to which there was disparity in the quantity of written feedback across classes. Students in the “academically gifted” classes received written feedback on more essay assignments and more written corrections or comments on their work compared to their peers in the “regular” classes. They also appeared to get more immediate feedback on their written work compared to students in the “regular” classes. This finding underscores the importance of researchers analyzing student work in addition to conducting extended classroom observations to uncover subtle classroom practices that contribute to students’ schooling experiences. In this final difference in instructional opportunity across tracks, the context of high-stakes accountability programs did not appear to influence these teachers to be aware of differences in the quantity of feedback across tracks, despite the state’s focus on the achievement gap.

IMPLICATIONS

Based on teacher testimony and classroom observations after implementing the high-stakes accountability program, the high-stakes state accountability policy, in large part, appears not to have changed the dynamic identified by previous researchers in regards to the disparity in the
quality of instruction across tracks (Oakes, 1985; Page, 1989). In the high-stakes statewide accountability context, students in the lower track had fewer opportunities to practice their reading and writing, less challenging instruction and less feedback compared to their peers in the higher track. While Oakes and Page did not write specifically about the nature of feedback across tracks, my findings are consistent with the disparities these researchers located across tracks. That is, students in the higher track were expected to do more reading and writing assignments through “a great deal of expository writing” and more homework than their peers in the lower track (Oakes, 1985, p. 76; Page, 1989). In addition, teachers emphasized simple comprehension for students in the lower track while they “demanded critical thinking, problem solving, drawing conclusions, making generalizations or evaluations or synthesizing knowledge” in the higher track (Oakes, 1985, p. 76). The high-stakes statewide accountability policy did not disrupt teachers’ sequential theories of learning, which influenced teachers to make these bifurcated curricular decisions for their “regular” and “academically gifted” students, nor did it address the state’s tracking policy, which provided the structure for teachers to enact their curricular decisions.

On the other hand, teacher testimony about instruction before and after high-stakes testing, as well as comments about what they would do differently without high-stakes testing, suggest that testing has influenced teachers to do more isolated reading skill instruction for all students, but particularly those in the “regular” track, because of the emphasis on explicit test preparation. Without testing, many teachers stated that they would have more time to pursue priorities they deemed important for students: to practice a variety of writing genres for “authentic” audiences, to complete more independent assignments where students worked on their communication and collaboration skills and where they developed an appreciation for literature (Watanabe, 2007, a). In the testing climate, however, explicit test preparation trumped these priorities, particularly for students in the “regular” track.

The student makeup of the two track levels in focal teachers’ classrooms further adds a troubling dimension to this finding, as students of color and students from low socioeconomic backgrounds are disproportionately represented in the “regular” track. Students of color and students from low socioeconomic backgrounds in the “regular” track are thus shortchanged in opportunities to learn even in the new context of statewide accountability policies that policymakers purport to be part of the state’s response to “close the achievement gap.”

These differences in instructional opportunity have far-reaching implications for students’ educational and career trajectories and their future
roles as citizens who exercise their right to vote. Students in higher-tracked classes were encouraged to do more independent work that required “critical thinking, self-direction and creativity” (Oakes, 1985, p. 84). Students in “academically gifted” classes also had more opportunities to engage in current events. These classroom activities equip the “academically gifted” students, who are disproportionately students from more economically and racially privileged backgrounds, with high-status knowledge and skills that prepare them to go to college, take positions of leadership in the workplace and to participate actively as informed citizens in our democracy. Meanwhile, students in the “regular” track do not have equal access to this knowledge and skill set.

The data presented in this article help illuminate one of the complex reasons why the achievement gap persists even in the new context of statewide accountability programs that policymakers purport to be part of the state’s response to “close the achievement gap”—the educational opportunities afforded students by track level are unequal, privileging students in the “academically gifted” track and shortchanging students in the “regular” track. If tracking persists, high-stakes accountability programs are unlikely to lead to the desired closure of the achievement gap. Therefore, this research suggests the need for a two-tiered response: one that focuses attention at the policy level on the deleterious effects of tracking, and one that supports teachers at the school and classroom level through professional development. More attention must be targeted to address inequitable classroom opportunities across tracks, and to support teacher inquiry into closing the achievement gap.

ADDRESS STATE TRACKING POLICIES AND INEQUITABLE CLASSROOM OPPORTUNITIES ACROSS TRACKS

Policymakers need to address the state’s tracking policy as part of their plan to close the achievement gap. Merely focusing attention on the performance of student subgroups, as No Child Left Behind stipulates, without specifically looking into equitable classroom practices, will likely do little to challenge teachers to think critically about the types of instructional opportunities they afford students. Reorganizing schools so that classrooms are detracked is one avenue to bring critical questions about equity to the forefront and to equalize students’ schooling experiences.

Skeptics of detracking initiatives may question why it is necessary to dismantle tracking. They may instead insist that schools need to more closely monitor the placement practices to ensure their fairness and that they need to improve the quality of instruction across tracks. Although these strategies may certainly improve the troubling patterns
documented in the research literature, the hierarchical nature of grouping practices always privileges one group of students over another, and “these groups are labeled quite openly and characterized in the minds of teachers and others as being of a certain type—high ability, low achieving, slow, average, and so on” (Oakes, 1985, p. 3). For example, in my research, Ms. Jones one day notes incredulously to students in her “regular” class that some students, and “this was AG people,” whined that they could not locate a simile in the text (9/13/01 SJ field notes, p. 10). Even though Ms. Jones makes this comment to poke fun at students in the “academically gifted” class, she still communicates to students in the “regular” class that she expects more from “academically gifted” students than “regular” students.

Students internalize identities of superiority/inferiority attached to these hierarchical labels. For example, during one of my observations, one “academically gifted” student uses the argument, “And we’re the best class. We’re AG,” to encourage Ms. Royster to stop teaching before the bell rings so students will have enough time to pack their bags and get to their next class period on time (11/6/01 MJR field notes, p. 11). In addition, when a “regular” student earns a 100 on a test in a social studies class, he says, “100? Oh! Oh my God, I guessed. Hey! Can I be in AG today?” (12/3/01 AJ field notes, p. 3). Even with a more proportional racial composition of students across tracks, tracking is problematic as educators treat hierarchical groups differently and students internalize and meet teacher and peer expectations for student performance (Dorph, 1999; Rosenthal & Jacobson, 1968; Rubin, 2001; Weinstein, 2002).

In addition, the state’s tracking policy provides the structure to allow teachers to unquestioningly enact their sequential theories of learning, when a large body of research suggests that isolated skill instruction and a focus on basic skills often perpetuates low reading skills instead of improving students’ facility with the complex process of reading (Allington, 2001; Alvermann & Moore, 1991; Brown, 1991; Hiebert, 1991; Hull & Rose, 1989; Knapp & Turnbull, 1991; Sizer, 1992, as cited in Commission on Reading, NCTE, 2004). The state’s tracking policy implicitly sets up the expectation that students in the “academically gifted” track should be engaged in more challenging curricula when all students should take part in meaningful, high-quality curricula that demands critical thinking.

Detracking, however, is not an easy undertaking, and instituting detracking as state policy may appear to be a daunting possibility for policymakers. Detracking has met with strong resistance, often from concerned parents of students who were previously in the higher tracks, who
fear that mixed-ability groups will hurt their child’s education (Wells & Serna, 1996). While I do not wish to minimize the complexity of detracking, a growing number of districts and schools have improved student achievement for all students, including previously high-tracked students, through detracking efforts, and policymakers could turn to these examples to support detracking policies with empirical evidence of their success (see Boaler & Staples, this issue; Burris, Welner, Wiley & Murphy, this issue; Wheelock, 1992). In addition, state policymakers implementing detracking policy would be joining other politicians, state education agencies, educator groups and other organizations working with youth, who issued policy statements condemning tracking. The National Governors Association, California and Massachusetts state departments of education, the National Education Association, National Council for the Social Studies, the National Council of Teachers of English, the National Council of Mathematics Educators, the NAACP Legal Defense Fund, the ACLU, the Children’s Defense Fund, the Carnegie Corporation Council for Adolescent Development, and the College Board, among others, publicly denounced tracking policies as harmful to students in the late 1980s and early 1990s (see Oakes, 2005).

Detracking, however, does not automatically disrupt institutional patterns of inequality at the classroom level (Cohen, Kepner, & Swanson, 1995; Cone, 1992; Rubin, 2003). Rubin (2003) illuminates the complex weave of factors that reproduce inequalities, even in detracked classroom settings, factors such as teacher and student assumptions about race, class, gender and academic competence, students’ desires to work in small groups with friends for fun and mutual respect, and the activation of out-of-school economic and cultural capital in classroom activities. At the same time, Rubin (2001) documents how detracking has the potential to transform patterns of social reproduction, to foster complex discussions where race is foregrounded, thereby enriching the classroom experiences for all students. These experiences are possible in detracked classroom settings when teachers attend to student placement in daily classroom activities, build community, create multiple points of entry, value varied forms of cultural capital, provide structure and explicit support for reaching higher expectations, and foster critical awareness on societal inequalities (Rubin, 2001).

While teachers may not be currently well equipped to think about ways to challenge students with different skill levels in the same classroom, research literature documents some successful strategies that serve as useful starting points (Cohen et al., 1995; Cone, 1990; Fine, Weis, & Powell, 1997; Horn, 2006; Oakes, Wells, & Associates, 1996; Rubin, 2001; Wheelock, 1992). In addition, models of teaching reading and writing,
such as reading apprenticeship, serve as alternatives to remedial approaches that can be used with heterogeneous groups of students (Schoenbach, Greenleaf, Cziko, & Hurwitz, 1999). These models disrupt the erroneous assumption that one must learn the “basics” sequentially in order to improve literacy skills.

In sum, my research provides further support to previous research literature on how membership in a particular track significantly influences the instructional opportunity a student receives even in the new school reform context of high stakes statewide accountability programs (Gamoran & Mare, 1989; Kerckhoff, 1986; Oakes, 1985; Page, 1991; Rosenbaum, 1978; Schafer et al., 1972). Given the disproportionate representation of Black students in the “regular” track and White students in the “academically gifted” track and the instructional opportunities afforded to students at each track level, tracking has the potential to perpetuate the achievement gap. With the proper support, detracking is one avenue to challenge this pattern.

SUPPORT TEACHER INQUIRY INTO CLOSING THE ACHIEVEMENT GAP

In addition, policymakers need to provide substantial support to teachers to initiate and sustain inquiry into explanations for the achievement gap that are both outside of and within their control (McLaughlin, 1987). Darling-Hammond (1998) emphasizes the importance of professional development that is site-based, long range and focused in contrast to one-shot professional development days that are often imposed by districts rather than generated from the school site. Teacher inquiry groups are one such model of professional development (Cochran-Smith & Lytle, 1999; Grossman, Wineburg, & Woolworth, 2001; Stokes, 1999; Wilson & Berne, 1999) that could potentially focus on issues of the achievement gap and the normative, technical and political dimensions of detracking (Oakes, 1992).

What might this recommendation look like in practice? I could envision a year-long school focus on the achievement gap, where teachers read about and reflect on the multifaceted explanations of the achievement gap, acknowledging factors that are outside of their realm of control and considering teaching and school practices that they can change (see Stokes, 1999). Teachers could read and discuss how social reproduction theories play out even in classrooms of well-intentioned teachers who care about their students, and follow up by asking themselves whether they are being attentive to all of their students. Furthermore, teachers could examine their theories of ability and intelligence, current
research on tracking and detracking and explore teaching methods for detracked classrooms (Watanabe, 2007a, b). Preservice teacher education programs could take leadership in beginning these conversations as well (Lotan, 2006).

While I do not want to reduce the complexity in the ways in which these professional development opportunities can unfold depending on the composition of the groups, teacher relationships, leadership, and ties to larger school structures that research on professional development documents (Curry, 2003; Grossman et al., 2001; Westheimer, 1998), teacher inquiry groups have gained increased visibility as promising professional development for teacher learning and school reform, deserving further attention as one model for policymakers and school administrators to consider.

Conversations about race and the roles we play in contributing to the achievement gap are no less difficult for researchers in relation to teachers than they are for teachers among themselves (Stokes, 1999). For example, Ms. Jones and Ms. Royster did not seem troubled by the preliminary findings I shared with them that students in the “academically gifted” track had more reading and writing opportunities, more challenging assignments and received more feedback than their peers in the “regular” track. These teachers each had explanations for these differences that made sense to them. Due to my own discomfort with how to go about challenging their explanations in a meaningful and respectful way, I refrained from voicing my concerns about equity issues during these interviews. This case represents a broader problem where dedicated teachers come up with explanations that do not challenge the achievement gap coupled with a research community looking closely at practice, who uncover areas that deserve attention but are unsure as to how to engage these conversations. As the educational and career trajectories of many students are at stake, I am convinced of the importance of learning to broach and documenting these complex and difficult discussions in many communities—among teachers and between researchers and teachers.

These difficult conversations have the potential to create “aha” moments for teachers that could promote powerful changes. For example, Stokes (1999) describes a teacher who realizes, through an action research project, that she only lends out books to students whom she assumes have parents who will ensure their return. She was therefore “unwittingly limiting the education of the most disadvantaged children” (p. 14). This teacher reflected on how this one example could be just the “tip of the iceberg” in the ways she afforded inequitable instructional opportunities to her students (p. 14). Such realizations can enable
teachers to gain a more nuanced understanding of how they contribute to inequities in students’ schooling experiences.

In sum, there is a critical need for educators to examine various explanations for the achievement gap, including their own roles in contributing to it. This reflection will better equip educators to think about ways to challenge the patterns of inequity that I document in my research. At the same time, merely adding professional development without adequate support (e.g., release time), would create a situation where policymakers would “add to (teachers’) plates, add to (teachers’) plates, add to (teachers’) plates” and lead teachers feeling “rattled” (SJ, Interview 4, p. 18). Policymakers must match support to teachers’ current work demands, through class size reduction and adequate planning and grading time during school hours. In addition, it is important for policymakers and administrators to create the time, space and support for teachers to engage in sustained and meaningful professional development.

CONCLUSION

The systemic inequalities in our nation’s schools are deeply troubling. These case studies demonstrate how North Carolina’s high-stakes state accountability policy fails to overcome structural inequalities already embedded in its schooling system, and in some cases, pushes teachers to respond in the direction of greater inequalities. Despite egalitarian claims, North Carolina’s high-stakes accountability policy does not overcome the stratifications of access to knowledge built into the practice of tracking. Moreover, some teachers are responding to the accountability policy in ways that widen the achievement gap; teachers spend more time on test practice with students in the “regular” track, predominantly Black students in this district, than students in the “academically gifted” classes, students more likely to be White and economically privileged. Completing isolated test preparation exercises for hours of class time instead of engaging in more substantive work which emphasizes critical thinking skills further widens the gap rather than closing it.

The state’s tracking policy is so embedded in the organization of schools that most of the teacher participants in this research accepted tracking as given. Teachers did not question the deleterious effects of tracking, or their sequential theories of learning that manifested in differential learning opportunities across tracks. However, teachers’ actions are shaped by the structural organization of schools and the underlying expectations these policies connote, even if teachers do not always explicitly identify these factors in discussions about their curricular decisions. Policymakers must thus dismantle policies that contribute to systemic
inequalities in our nation’s schools in addition to encouraging teachers to reflect on their teaching practices at the classroom level through professional development.

Author Note

This project was funded in part by a Spencer Center for the Integrated Study of Teaching and Learning grant and a Humanities Research grant at the University of California, Berkeley. I would like to thank Becky Cox, Marnie Curry, Linda McNeil, Pedro Noguera and Jeannie Oakes for their feedback on earlier drafts of this article.

Notes

1 Policymakers have added a new dimension to school reforms of the past to create high-stakes statewide accountability programs. From minimum-competency testing (1970s) and the alignment of testing to state standard frameworks (mid-1980s), policymakers in the 1990s have attached high stakes to standardized tests for schools, teachers, principals, and students in the form of punitive sanctions or performance rewards (Linn, 2000). Decisions about students’ grade promotion, principal promotion/demotion and cash bonuses for teachers and schools make current accountability reforms “high stakes.”


3 For example, the Evaluation Section of the Division of Accountability in the North Carolina Department of Public Instruction (2000) conducted a study of nine schools that demonstrated success in raising student achievement and in reducing the achievement gap since 1994. State bureaucrats identified the following common themes across school sites that they believed contributed to the high levels of achievement: leadership, focused instruction (e.g., pacing guides, curriculum integration, lesson plan alignment checks, professional development), periodic testing, disaggregating data, targeted intervention (e.g., technology for at-risk students, tutoring, small-group programs, class size reduction), a culture of achievement (e.g., “academic pep rallies”). The accountability program, and specifically attention to testing, emerges as a strong theme in these schools’ responses to promote high student achievement and to close the achievement gap.

4 School, teacher and district names are pseudonyms to protect the confidentiality of the people involved.

5 Policymakers added two new performance categories, schools of progress and priority schools, in reporting the 2001-2002 school year results. Policymakers did not attach monetary rewards to these performance categories. Schools of progress were schools that had met or exceeded state expectations for growth and had between 60% and 79% of students performing at or above grade level. Priority schools were a) schools with between 50 to 59% of students performing at grade level and b) schools with less than 50% of students at grade level but which had met or exceeded the growth standard (North Carolina
Department of Public Instruction, 2002, p. 2). The “no recognition” school classification was changed to identify a school that had not met state expectations for growth but had over 60% of students performing at grade level.

6 The exact percentage of White students performing at or above grade level was not reported; instead, the North Carolina Department of Public Instruction reported the percentage as greater than or equal to 95% (North Carolina Department of Public Instruction, 2005).

7 One exception was the quantity of feedback, for which insufficient data limited any cross-school analyses, since I did not analyze work from schools aside from Sequoia.

8 In a sociocultural approach to literacy, learning emerges in the context of social interaction (Vygotsky, 1978) and students are asked to draw upon their cultural background to construct meaning (Cazden, 1988; Phillips, 1972).

References


Testing Section, NC Department of Public Instruction (1999). *Understanding testing at Grades 3–8: Background,* 6(1), 1–2.


MAIKA WATANABE is Assistant Professor in the Department of Secondary Education at San Francisco State University. Broadly speaking, her teaching and research interests focus on the sociology of education, urban school reform with a focus on equity, and educational policy. Two recent publications include: “Displaced teacher and state priorities in a high-stakes accountability context” in *Educational Policy*, and “Lessons from a teacher inquiry group about tracking: Perceived student choice in course-taking and its implications for detracking reform” in *Teachers College Record*. 