

JOURNAL OF
MATHEMATICS
EDUCATION
AT TEACHERS COLLEGE

A Century of Leadership in Mathematics and Its Teaching

Beyond Teaching Mathematics

© Copyright 2016 by the Program in Mathematics and Education

TEACHERS COLLEGE | COLUMBIA UNIVERSITY

TABLE OF CONTENTS

PREFACE

- v *William McGuffey, Teachers College, Columbia University*
Kimberly Barba, Teachers College, Columbia University

ARTICLES

- 1 **Some Thoughts on Educating More Able Students**
Geoffrey Howson, University of Southampton, England
- 7 **A Comparison of Mathematics Teachers' and Professors' Views on Secondary Preparation for Tertiary Calculus**
Carol Wade, The College at Brockport, State University of New York; Gerhard Sonnert, Harvard University; Philip Sadler, Harvard University and the Smithsonian Institute; Zahra Hazari, Florida International University; Charity Watson, Clemson University
- 17 **The Legacy Continues: "The Test" and Denying Access to a Challenging Mathematics Education for Historically Marginalized Students**
Richard Kitchen, University of Denver; Sarah Anderson Ridder, University of Denver; Joseph Bolz, University of Denver
- 27 **Teaching Mathematics for Social Justice: Examining Preservice Teachers' Conceptions**
Cindy Jong, University of Kentucky
Christa Jackson, Iowa State University
- 35 **Fostering Justification: A Case Study of Preservice Teachers, Proof-Related Tasks, and Manipulatives**
Jonathan D. Bostic, Bowling Green State University
- 45 **Playing with Mathematics: How Play Supports Learning and the Common Core State Standards**
Jennifer Mary Zosh, Pennsylvania State University, Brandywine; Brenna Hassinger-Das, Temple University; Tamara Spiewak Toub, Temple University; Kathy Hirsh-Pasek, Temple University; Roberta Golinkoff, University of Delaware
- 51 **Assessment of a Problem Posing Task in a Jamaican Grade Four Mathematics Classroom**
Kayan Lloyd Munroe, Hiroshima University

The Legacy Continues: “The Test” and Denying Access to a Challenging Mathematics Education for Historically Marginalized Students

Richard Kitchen
University of Denver

Sarah Anderson Ridder
University of Denver

Joseph Bolz
University of Denver

ABSTRACT Research is needed to understand the impact of high-stakes testing on teachers’ practices and consequently on their students, particularly at schools that serve large numbers of low-income students and students of color. In this research study, we examined how a state’s annual high-stakes test and administrative mandates influenced the assessment practices of mathematics teachers at a highly diverse urban high school. Among the most compelling findings was that students at Chavez High (a pseudonym) were labeled based upon their test performance, that this label tended to persist at the school, and that instructional decisions were made based upon these labels. This practice of attaching a label to students based upon their performance on the high-stakes test constructed some students as less capable than others in mathematics. Such practices contribute to the historic legacy in the U.S. of denying poor students and students of color access to a challenging education in mathematics.

KEYWORDS *mathematics at diverse schools, assessment, test prep in mathematics*

Introduction

NCLB was a punitive law based on erroneous assumptions about how to improve schools... Perhaps most naively, it [NCLB] assumed that higher test scores on standardized tests of basic skills are synonymous with good education. Its assumptions were wrong. Testing is not a substitute for curriculum and instruction. Good education cannot be achieved by a strategy of testing children, shaming educators, and closing schools. (Ravitch, 2010, p. 110-11)

In the quotation above, Diane Ravitch, well-known critic of school privatization and the No Child Left Behind Act (2001), problematizes the role that high-stakes testing is playing in the United States (U.S.) and the impact that it is having on what teachers teach and how they teach it. High-stakes testing involves “attaching impor-

tant consequences to standardized test scores, and it is the engine that drives the No Child Left Behind (NCLB) Act” (Nichols & Berliner, 2008, p. 672). Many others have made similar arguments about the detrimental effect of the NCLB legislation (2001) to prepare students for high-stakes testing and how the work of teachers is reduced to focusing on the instruction of low-level skills that are tested (see, for example, Burch, 2009; Valenzuela, 2005). Interestingly, the role and limitations of high-stakes tests in terms of what they are designed to accomplish were well articulated by the assessment community prior to the NCLB legislation (see, for example, Linn, 2000; Messick, 1995a; 1995b), but were largely ignored by policy-makers.

At a time when high-stakes testing, such as No Child Left Behind mandated tests (NCLB, 2001) have dominated the educational landscape in the U.S., more research is needed to understand the impact of these tests not just on teachers’ practices, but on students, particularly at

schools that serve large numbers of low-income students and students of color since these students have historically been underserved by public education (Darling-Hammond, 1996; Ferguson, 1998; Knapp & Woolverton, 1995; Massey, 2009; Milner, 2013). The purpose of this research study is to examine how the assessment practices of mathematics teachers at a highly diverse urban high school in the U.S. were influenced by their state's annual high-stakes test and how these practices impacted low-income students and students of color.

Standardized Tests

Supporters of high-stakes testing believe that the quality of American education can be vastly improved by using high-stakes assessments to introduce a system of rewards and sanctions that are triggered by students' standardized test performance (Raymond & Hanushek, 2003). The theory of action undergirding this approach is that educators and their students will work harder and more effectively to enhance student learning when faced with large incentives and threatening punishments. Some researchers argue, however, that serious problems accompany the introduction of high-stakes testing.

Standardized tests typically assess lower-order thinking skills through the use of easily administered items (Newmann, Bryk, & Nagaoka, 2001). Since teachers are reluctant to embrace new assessment practices unless these practices align with high-stakes testing, it is not surprising that teachers' assessment practices may support traditional, lower-order mathematics instruction (Barnes, Clarke, & Stephens, 2000; Nichols, Glass, & Berliner, 2006; Wilson, & Kenney, 2003). Critics also worry that the pressures to do well on a test may lead to "corrupt" educational practices (Nichols & Berliner, 2007). Categorizing students by their achievement on standardized tests (for instance, being labeled "below basic") can also "reinforce feelings of marginalization that already impact the achievement of many students of color and others ill-served in schools" (Duckor & Perlestein, 2014, p. 27).

"Teaching to the test" reduces students to test performers and "teachers find themselves using students to protect or help themselves.... The marketplace mentality expands its reach" (Gergen & Dixon-Román, 2014, p. 8). From this perspective, testing plays a powerful role in controlling teachers' work (Apple, 2014). As the evaluation of teachers continues to expand to take into consideration students' performance on standardized tests, teachers feel more compelled to prepare their students for success on the test just to be able to keep their jobs

(Apple, 2014). Standardized testing also essentializes students into a test score or a set of test scores that are then used by those in power (administrators and teachers) to make decisions that impact students' lives (Gergen & Dixon-Román, 2014). Finally, instruction may be more targeted to those students who are close to meeting proficiency on the test, and less instructional attention is paid to those who are far above or below meeting proficiency (Harlen, 2007).

High-stakes tests cannot meet all the demands made on them by states and school districts (Linn, 2000; Messick, 1995a, 1995b). For instance, in standardized testing, priority is placed on the psychometric techniques for producing valid and reliable standardized scores. Moreover, many significant challenges are associated with designing standardized assessments aligned with the ideals of mathematics education that meet standards of measurement theory (Graue, 1993; Shepard, 2000; Suurtamm, Lawson, & Koch, 2008). Shepard (2000) and Pegg (2003) call for changes in assessment to reflect a change in practice in which students actively make meaning of mathematical concepts by building on their previous knowledge. Specifically, assessment formats should align with reforms proposed in mathematics (NCTM, 1989, 1995, 2000; NSF, 1996) in which the primary goal is for students to develop mathematical understanding through problem solving by making connections, communicating, and representing mathematical ideas (Gamoran, Anderson, Quiroz, Secada, Williams, & Ashman, 2003). Because of the focus on preparing students for success on high-stakes tests, assessments that include rich mathematical problems for use at the classroom-level have been receiving little attention in the US (Kulm, 2013).

Mathematics Education of Low-Income Students of Color

There is a need for more research that specifically addresses how to improve the mathematics education of students of color (also referred to as *diverse students* throughout; students who are members of a racial or ethnic minority group. For us, *diverse students* is synonymous with *culturally and linguistically diverse students* and *students of color*) and students from poverty precisely because these students have historically been denied access to a challenging mathematics education in the US (Davis & Martin, 2008; Kitchen, DePree, Celedón-Pattichis, & Brinkerhoff, 2007; Martin & Leonard, 2013; Téllez, Moschkovich, & Civil, 2011). The historic legacy of classism in the US, and racism against Blacks, Native Americans, and Latinos/as, in particular, provide the

backdrop for the impoverished educational system that we find today in urban and highly rural districts that primarily serve low-income students and students of color (Darling-Hammond, 1996; Ferguson, 1998; Knapp & Woolverton, 1995; Massey, 2009; Milner, 2013). In mathematics, students have historically been sorted and stratified (e.g., by tracking) by race, ethnicity, class, and gender (Diversity in Mathematics Education [DiME] Center for Learning and Teaching, 2007; Gerdes, 1988). Specifically, White and Asian middle class and upper-middle class students have been privileged to have greater access to challenging mathematics curriculum and instruction (DiME, 2007; Tate, 1995). Schools that enroll large numbers of African American students often have disproportionately high numbers of remedial classes in mathematics in which instruction is focused on rote-learning and strategies that are intended to help students be successful on standardized tests (Davis & Martin, 2008; Lattimore, 2005; Martin, 2013). In schools that serve large numbers of immigrant Latino/a students who speak with an accent, use English words incorrectly or speak in Spanish as a means to express themselves, educators, peers and community members may assume they lack the capacity to perform well in mathematics (Gutiérrez, 2008; Moll & Ruiz, 2002; Moschkovich, 2007).

Ability grouping or tracking is a commonly used practice in the US that has disproportionately hurt the poor and students of color (Oakes, 2005; Secada, 1992). Ability grouping, or tracking, involves grouping students in the classroom according to their talents (Oakes, 2005). Proponents of tracking believe that the practice allows teachers to tailor the pace and content of instruction to students' needs (<http://www.nea.org/tools/16899.htm>). Opponents, however, contend that tracking not only fails to benefit any student, but it also channels poor students and students of color to low tracks where they receive a lower quality of instruction than other groups. Tracking continues to "divide students by perceptions of 'ability' and communicate to students the idea that only some people—particularly white, middle class people—can be good at mathematics" (Boaler, 2011, p. 7). In general, educators of racial/ethnic minorities often make the memorization of mathematical facts, algorithms, vocabulary, and procedures the focal point of their instruction, rather than teaching students through the use of complex, challenging problems (Davis & Martin, 2008; Kitchen et al., 2007; Lattimore, 2005). Historically, millions of low-income students and students of color have been denied access to a challenging, problem-solving based curriculum provided to their counterparts in more affluent communities (Davis & Martin, 2008; Flores,

2008; Kitchen, Burr, & Castellón, 2010; Payne & Biddle, 1999). Generally speaking, the power elite in the US (e.g., White and middle- and upper-class Americans) have historically not opposed high-stakes testing because they knew their children were prepared to perform well on the tests, thus preserving their advantages (Nichols & Berliner, 2007).

Methodology and Data Sources

We purposefully conducted this study at Chavez High School (a pseudonym), a highly diverse urban high school in Colorado. By highly diverse, we mean that a high percentage of "diverse students" attended the school. The study was carried out in 2013-2014 with five Chavez High mathematics teachers. At the time, Chavez High served slightly more than 2,000 students in grades 9-12. In 2013-2014, the Free Reduced Lunch rate at Chavez High was 75%. Sixty-four percent of the Chavez High student body was Hispanic, 15% identified as African American, 13% was White, 4% identified as Asian American, 3% were two or more races, and 1% was Native American or Native Hawaiian. In 2013-14, more than 60% of the school's students spoke a language other than English at home.

Though Colorado is a PARCC state (The Partnership for Assessment of Readiness for College and Careers, or PARCC, is a standardized test that was initially administered in Colorado and other participating states in spring 2015), the standards-based Transitional Colorado Assessment Program (TCAP) assessment in mathematics was administered to all students in grades 3-10 in the state in the spring of the 2013-2014 school year. The research question that we address in this paper is: How did the state mandated high-stakes test in mathematics and administrative mandates influence the assessment practices of mathematics teachers at Chavez High?

The data collected to answer the research question included classroom artifacts used during lessons observed such as student work samples (e.g., worksheets completed, problem solutions, etc.), notes taken during classroom observations, videotapes recorded during observed lessons, and audiotaped interviews. Individual interviews were conducted with each participating teacher and with the Chavez High principal. We also conducted focus group interviews with the participating teachers as a whole.

Five mathematics teachers at Chavez High provided consent to participate in this study as did the school's principal. Beginning in the fall semester of 2013, we vis-

ited Chavez High 2-4 times per month for the duration of the 2013-2014 school year. A school visit included a classroom observation of a participating mathematics teacher on consecutive days. A classroom observation consisted of videotaping the participating teacher teaching a mathematics lesson as well as videotaping a group or groups of students who had provided consent to participate in the study. Every attempt was made to videotape in a manner that minimized interference in the mathematics lesson (e.g., the video camera was placed in a location in the classroom such as the back of the room so as not to block students' view of their teacher, the whiteboard, and any other instructional resources used by the teacher). An interview was conducted immediately following the first or second classroom observation. Interviews with individual teachers were 30-45 minutes in length and were videotaped. Three focus group interviews were also conducted with the teachers: two in spring 2014 and a third that took place exactly one year following the second focus group interview in spring 2015. Four of the five participating teachers participated in the initial focus group interview, while all five of the participating teachers attended both the second and third focus group interviews. An interview was conducted with the Chavez High principal in spring 2014 as well. The lead author conducted all the interviews with assistance, on occasion, from the second and third authors.

The data analyzed included all the notes made by researchers during classroom observations and information collected during the teachers' Professional Learning Community (PLC) meetings, and the videotaped interviews were conducted with participating teachers and the principal. The interview transcripts were analyzed using interpretive methods (Erickson, 1986; Maxwell, 2005). Each interview was read as a whole, followed by a period of open coding to allow for the emergence of themes, and themes were then compared across interviews conducted. After a set of themes was obtained from the dataset, we searched for commonalities and differences across interviews conducted (Miles, Huberman, & Saldaña, 2013). We also sought both confirming and disconfirming evidence by searching for supportive and non-supportive evidence (Erickson, 1986; Miles, Huberman, & Saldaña, 2013).

The five mathematics teachers at Chavez High who participated in this study were given the pseudonyms Ms. A, Ms. B, Ms. C, Mr. D, and Ms. E. All five teachers were chosen by the school's administration as among the best mathematics teachers at the school, were recom-

mended for inclusion in this study, and provided consent to participate in this study. Ms. A and Ms. B were the most veteran teachers at Chavez High, both having served at the school for 10 or more years. At the time this study was undertaken, Ms. A was the chair of the Mathematics Department at Chavez High. Ms. C is Hispanic and was the only teacher of color in the group of participating teachers. Mr. P, the Chavez High principal, was also a participant in this study.

How "The Test" Impacted Teachers' Assessment Practices

We found that Colorado's high-stakes test ("The State Test") impacted teachers' assessment practices in at least three ways. First, we learned that the language used to characterize student performance on The State Test became normalized by the teachers at Chavez High. Second, we found that the teachers used the language used to characterize student performance on The State Test in the design of assessments. Finally, we learned that the language used to characterize student performance on The State Test was also used to assign students their grades in the mathematics classes at the school.

Testing Language as Normalized Through Assessment

Because of the intense focus at Chavez High on students' performances on The State Test, it was common for teachers and administrators to refer to students as Advanced, Proficient, Partially Proficient or "Bubble Students," or as Unsatisfactory (or just as "Unsats"). This language reflected not only how students had performed on The State Test, but had also become language that teachers had adopted to design assessments and even assign students grades. Language commonly used to discuss assessments and students at Chavez High not only reflected the language used on The State Test, but phrases such as "low hanging fruit" had been adopted as a way to delineate which students could essentially be reached and had a chance to be proficient and those that did not, "the Unsats."

It was common to hear teachers and administrators refer to a specific student as a "Bubble Student" or as an "Unsat student." In interviews conducted with teachers and the school principal and in conversations that our research team observed that took place during PLCs, we learned that it was a common practice for teachers and school administrators to refer to students by these labels. We also found the "Bubble Student" or "Unsat" label

persisted during the year we spent at Chave High; it was well documented who these students were and teachers knew who, among their students, was a Bubble Student and who was an Unsat student. Teachers were encouraged to devote what limited time they had to work just with those students who had shown potential to score Proficient or higher on the next administration of The State Test. Bubble students were specifically targeted for supplemental instruction since these students were within reach of achieving Proficient on The State Test.

Testing Language Used to Label Student Performance and Design Assessments

Though students were classified as a Bubble Student or as Unsat based on their performance on the previous year's State Test, teachers also used the testing categories to label student performance on classroom assessments. For instance, mathematics teachers at Chavez High engaged in what was referred to as "deployment," another assessment practice at the school that was mandated by the administration. Students who were evaluated as Partially Proficient or Unsatisfactory on a unit test were "deployed" to receive supplementary instruction on the mathematics unit just completed. After the 1-2 day deployment, students were administered a post-test that included mathematical content that was similar to what was included on the original unit test. When discussing the impact of one such deployment, Ms. B reported that "77% of the Unsat's moved up to Partially Proficient" on a post-test administered.

Teachers also discussed elements of assessments they designed using the classification language used on The State Test. For instance, parts of unit tests utilized for deployments included "Unsat items," or questions that generally required less of students (e.g., recall of mathematical vocabulary). Ms. C noted in a focus group interview that the "Unsat portion" of an assessment was meant to provide "access points" for students. By this, Ms. C meant that the content of the Unsat portion should be more elementary, focusing for example, on vocabulary items in geometry. Teachers explained that the majority of the tasks on assessments such as Exit Tickets were Unsat items (e.g., 70-80% of the items), while the remainder of the tasks on the assessment were more advanced. Mr. D explained that he generally included only "the naked" or skills based problems on his Exit Tickets. Here, we see that the teachers used the language of The State Test to refer to items designed for use on assessments, rather than classify items based on the grade a student was receiving (e.g., an A, B, C, D, or F student). The content of mathematics assessments was clearly in-

fluenced by the content teachers perceived would be included on The State Test. On many occasions, the teachers discussed constructing assessment items as a means to gauge student preparedness for The State Test.

Testing Categories Used for Grade Assignment

Mathematics teachers at Chavez High used The State Test categories to assign grades to their students. For instance, Ms. C used the classification language used on The State Test as part of her grading system, she did not assign grades based upon percentages: "I don't have any numbers in my grade book. It's a Partial, Unsat, or Advanced." For Ms. C, grades were assigned based upon how each student was performing, in a holistic manner, relative to language aligned with the test:

Unsat is they have some knowledge of some of the math that we did, so that's about a D. Partial Proficient is C-ish. And Proficient is about a B because you're doing what the standards are asking you to do. For students to earn an A, they have to take the math they've been doing and apply it to new problems that hasn't been taught to them.

In summary, we learned that Colorado's standards-based State Test profoundly impacted teachers' assessment practices. Because of the strong focus by the Chavez High administration on improving student achievement on The State Test, much of how teachers classified students and designed instruction was driven by the need to raise student scores on The State Test. Perhaps more importantly, student performance on The State Test had a strong influence on how teachers viewed students, and consequently, how they viewed them in the assessment process. The language that was used on The State Test to classify students by their performance (i.e., Advanced Proficient, Proficient, Partially Proficient, or Unsatisfactory) was used as a means to label students, even during the academic year following the administration of The State Test, and had become the taken-for-granted language teachers had adopted in their design of assessments and even in how they approached assigning grades to students.

Administrative Mandates and Policies Impacted Teachers' Assessment Practices

While studying how The State Test influenced the assessment practices of the mathematics teachers at Chavez High, we learned that the teachers' assessment practices were also affected by the school's administration. In

addition to their passion for improving student performance on The State Test, Chavez High administrators were under pressure to increase, in general, the number of students at the school who were passing their mathematics classes and enrolling in honors-level courses. In the following, we describe how mandates intended to limit student failure and increase the number of students enrolled in honors mathematics courses impacted the mathematics teachers at Chavez High.

Limiting Failure and Increasing Enrollment in Honors Classes at Chavez High

According to the teachers, there was an unwritten rule at Chavez High about the number of students that were allowed to fail in a given teacher's class. In a focus group interview, the teachers indicated that they were uncertain whether this rule was part of the school's actual policies, but they were certain that a failure rate of 20% or higher for a given class led to a meeting with a school administrator. As a response to this unwritten rule, Chavez High students were often allowed to take an assessment more than once to improve their grade. Teachers expressed frustration that students were often allowed to retake assessments on which they had not performed well. Ms. A discussed in a focus group interview a concern shared by her colleagues about how some students approached assessments. "I'm sick and tired of hearing that before I hand out the test, 'there is a re-test, right?'" Teachers believed that students were not being held accountable enough in the assessment process at Chavez High and were learning that they could always retake an assessment if their results were not good. Ms. A summarized the focus group's concern about this mandate in the following: "How is that kid ever going to learn how to succeed if they never actually get to fail?"

The teachers also described how the school's principal wanted as many students as possible to enroll in honors-level mathematics courses. However, the teachers did not believe that many of their students were ready to enroll in these classes. Ms. C shared the following in a focus group interview:

[The principal] pads his numbers in the sense that... 'Oh I have all of these honors classes....' The students aren't honors level and so when the teachers keep the classes at an actual honors level, it's our fault that the kids can't... that the kids aren't passing.... They recommended, I think it was two courses for honors for next

year for sophomores. We ended up with five. He's putting Unsat kids that have 3rd grade reading levels [in honors classes], just so he can say that he has five honors classes.

Students were also being allowed to skip classes in the progression of the integrated mathematics curriculum at the school. The integrated mathematics curriculum used at Chavez High was designed sequentially as a "spiraling curriculum" (i.e., topics are revisited throughout the curriculum, each time at successively more advanced levels). The curriculum was a three year program; Core I in year one, Core II the following year, and Core III in the third year. Students in honors classes took classes in algebra, trigonometry and calculus. In these classes, more "traditional" textbooks were used.

Mr. D noted in an interview that students were allowed to skip the Core II class because they had done well in Core I. In response, Ms. E said in a focus group interview, "No, they were like D and F students, and they won't let us bring them back down." There was much agreement among the teachers that the reason why students were being allowed to skip classes, even students who did not have good grades, was because of the principal's goal to increase the number of students who were enrolled in higher-level mathematics courses. From the teachers' perspective, this forced them to lower academic standards so that the requisite 80% of students were awarded passing grades in their mathematics classes.

Discussion

Few studies exist in the U.S. that specifically examine how low-income students of color are impacted by a state's high-stakes, standardized test. The participating mathematics teachers at Chavez High provided insights about how students' achievement on The State Test are also connected to other administrative mandates and policies at the school that directly influenced the teachers' assessment practices. Given the significant role that districts currently play in the U.S. to meet mandates of federal legislation such as NCLB, school administrators and teachers receive significant direction from the district central office (Datnow, 2012). In addition to the teachers' assessment practices at Chavez High being largely influenced by administrative mandates to improve student achievement in mathematics on The State Test, the teachers believed their assessment practices

were compromised to limit failure rates in mathematics classes at the school and increase the number of students who enrolled in honors mathematics courses.

School-change theorists argue that teachers' decision-making powers have been reduced (Hargreaves & Shirley, 2009), though teachers as active agents can engage in, passively accept, or reject central office mandates and policies (Olsen & Sexton, 2009; Payne, 2008). The teachers explained on many occasions that the administrative mandates and policies imposed on them were all intended to improve the image of Mr. P, the Chavez High principal. While Mr. P was under tremendous pressure from the district to raise test scores, increase the number of students enrolled in honors courses, and improve graduation rates at Chavez High, the teachers adamantly believed that Mr. P viewed the school's success in these areas as validating his leadership skills. In our interview with Mr. P and in casual conversations, he frequently made reference to the school's improved profile that he argued was largely based on increased test scores and more students taking advanced courses. Nevertheless, only about 20% of students at Chavez High scored Proficient or Advanced on The State Test administered in spring 2014.

Administrators and teachers work within a structure of power (Foucault, 1978; 1980) in which district mandates imposed on them impact their beliefs about students, instruction and assessment (Hargreaves & Shirley, 2009). The power that The State Test had taken on at Chavez High influenced how assessments were designed at the school and how grades were assigned to students. Moreover, similar to Gergen & Dixon-Román (2014), we found that students at Chavez High were labeled based upon their performance on The State Test, that this label tended to persist during the academic year at the school, and that instructional decisions were made based upon these labels. This practice of attaching a label to students based upon their performance on The State Test constructed some students as less capable than others in mathematics and some as more capable. The "Unsat" label persisted over the year at Chavez High and was difficult for students to shake. During the teachers' PLC meeting and in interviews conducted with them, they frequently talked about particular students based upon their performance on The State Test (e.g., Unsat students). Perhaps more importantly "Bubble students" were specifically earmarked for supplemental instruction given that they were close to scoring proficient on The State Test. The fact that students labeled as "Unsat" may not have had access to as much supplemental

instruction in mathematics as "Bubble students" is clearly problematic (Harlen, 2007). Given their low performance on The State Test, these students deserved more mathematics instruction, not less.

In 2014-2015, the first year at Chavez High that the PARCC exam was administered, all students at the school took two mathematics courses. Doubling up on mathematics courses was a strategy that the administration undertook to provide supplemental instruction to all students at Chavez High, not just the "Bubble students." A discussion of the varied impacts of this policy on teachers and students at the school is beyond the scope of this article.

Final Remarks

This study demonstrates that the response to both district and school-level administrative pressures to prepare students for success on the state's high-stakes test at Chavez High in 2013-14 contributed to the educational legacy of denying poor students and students of color access to a challenging mathematics education (Davis & Martin, 2008; DiME, 2007; Flores, 2008; Gutiérrez, 2008; Jacobsen, Mistele, & Sriraman, 2013; Kitchen, Burr, & Castellón, 2010; Martin, 2000; Martin, 2013). Specifically, attaching a label such as "Unsat" to any student is unjust—it is particularly unjust considering that the majority of Chavez High students originated from marginalized and oppressed communities that have historically not had access to a challenging mathematics education in the U.S. (Davis & Martin, 2008; Kitchen, DePree, Celedón-Pattichis, & Brinkerhoff, 2007; Martin & Leonard, 2013). At Chavez High, students were largely constructed as either capable or not capable in mathematics based upon their performance on the state's high-stakes test. Once students were mathematically classified relative to one of the performance categories on The State Test, supplemental instruction was allocated to those who were viewed as "low hanging fruit," students who were within reach of attaining proficiency on this test. Instructional interventions were created specifically to support "Bubble students," to increase the percentage of students at the school who scored proficient or better on The State Test. In the process, "Unsat" were denied opportunities to much needed supplemental instruction, given the belief of the Chavez High administration that resources were limited and that only so many students could attain proficiency on the state's high-stakes test.

References

- Apple, M. (2014). *Official knowledge: Democratic education in a conservative age* (3rd ed.). New York: Routledge.
- Barnes, M., Clarke, D., & Stephens, M. (2000). Assessment: The engine of systemic curricular reform? *Journal of Curriculum Studies*, 32(5), 623–650. doi: 10.1080/00220270050116923
- Boaler, J. (2011). Changing students' lives through the de-tracking of urban mathematics classrooms. *Journal of Urban Mathematics Education*, 4(1), 7–14.
- Burch, P. (2009). *Hidden markets: The new education privatization*. New York: Routledge.
- Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, and practice for democratic education. *Educational Researcher*, 25(6), 5–17.
- Datnow, A. (2012). Teacher agency in educational reform: Lessons from social networks research. *American Journal of Education*, 119(1), 193–201.
- Davis, J., & Martin, D. B., (2008). Racism, assessment, and instructional practices: Implications for mathematics teachers of African American students. *Journal of Urban Mathematics Education*, 1(1), 10–34.
- Diversity in Mathematics Education (DiME) Center for Learning and Teaching. (2007). Culture, race, power and mathematics education. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (405–433). Charlotte, NC: Information Age.
- Duckor, B., & Perlstein, D. (2014). Assessing habits of mind: Teaching to the test at Central Park East Secondary School. *Teachers College Record*, 116(2), 1–33.
- Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.) (119–161). New York: Macmillan.
- Ferguson, R. F. (1998). Teachers' perceptions and expectations and the Black-White test score gap. In C. Jencks & M. Phillips (Eds.), *The Black-White test score gap*, (273–317). Washington, DC: The Brookings Institution.
- Flores, A. (2008). The opportunity gap. In R. S. Kitchen & E. Silver (Eds.), *Promoting high participation and success in mathematics by Hispanic students: Examining opportunities and probing promising practices [A Research Monograph of TODOS: Mathematics for All]*, 1, 1–18. Washington, DC: National Education Association.
- Foucault, M. (1978). *The history of sexuality. Volume one: An introduction*. New York: Pantheon.
- Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972–1977*. New York: Pantheon.
- Gamoran, A., Anderson, C.W., Quiroz, P.A., Secada, W.G., Williams, T., & Ashman, S. (2003). *Transforming teaching in math and science: How schools and districts can support change*. New York: Teachers College Press.
- Gerdes, P. (1988). On culture, geometrical thinking and mathematics education. *Educational Studies in Mathematics*, 19(2), 137–162.
- Gergen, K. J., & Dixon-Román, E. J. (2014). Social epistemology and the pragmatics of assessment. *Teachers College Record*, 116(11). Retrieved at <http://www.tcrecord.org.libproxy.unm.edu/library/Issue.asp?volyear=2014&number=11&volume=116>
- Graue, M. E. (1993). Integrating theory and practice through instructional assessment. *Educational Assessment*, 1(4), 283–309.
- Gutiérrez, R. (2008). A “gap gazing” fetish in mathematics education? Problematising research on the achievement gap. *Journal for Research in Mathematics Education*, 39(4), 357–364.
- Hargreaves, A., & Shirley, D. (2009). *The fourth way: The inspiring future for educational change*. Thousand Oaks, CA: Sage.
- Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15–28.
- Jacobsen, L. J., Mistele, J., & Sriraman, B. (Eds.). (2013). *Mathematics teacher education in the public interest: Equity and social justice*. Charlotte, NC: Information Age Publishing.
- Kitchen, R. S., Burr, L., & Castellón, L. B. (2010). Cultivating a culturally affirming and empowering learning environment for Latino/a youth through formative assessment. In R. S. Kitchen & E. Silver (Eds.), *Assessing English language learners in mathematics* (Vol. 2) (59–82). Washington, DC: National Education Association.

- Kitchen, R. S., DePree, J., Celedón-Pattichis, S., & Brinkerhoff, J. (2007). *Mathematics education at highly effective schools that serve the poor: Strategies for change*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Knapp, M. S., & Woolverton, S. (1995). Social class and schooling. In J. Banks & C. Banks (Eds.), *Handbook of research on multicultural education* (548–569). New York: Macmillan.
- Kulm, G. (2013). Back to the future: Reclaiming effective mathematics assessment strategies. *Middle Grades Research Journal*, 8(2), 1–10.
- Lattimore, R. (2005). African American students' perceptions of their preparation for a high-stakes mathematics test. *The Negro Educational Review*, 56 (2 & 3), 135–146.
- Linn, R. L. (2000). Assessments and accountability. *Education Researcher*, 29(2), 4–15.
- Martin, D. (2000). *Mathematics success and failure among African-American youth: The roles of sociohistorical context, community forces, school influence, and individual agency*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Martin, D. B. (2013). Race, racial projects, and mathematics education. *Journal for Research in Mathematics Education*, 44(1), 316–333.
- Martin, D. B. & Leonard, J. (Eds.). (2013). *Beyond the numbers and toward new discourse: The brilliance of Black children in mathematics*. Charlotte, NC: Information Age Publishing.
- Massey, D. S. (2009). The age of extremes: Concentrated affluence and poverty in the twenty-first century. In H. P. Hynes & R. Lopez (Eds.), *Urban health: Readings in the social, built, and physical environments of U.S. cities* (5–36). Sudbury, MA: Jones and Bartlett Publishers.
- Maxwell, J. (2005). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications, Inc.
- Messick, S. L. (1995a). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice*, 14(4), 5–8.
- Messick, S. L. (1995b). Validity of psychological assessment: Validation of inferences from person's responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749.
- Miles, M.B., Huberman, A.M., & Saldaña, J. (2013). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: Sage Publications, Inc.
- Milner, H. R. (2013). Analyzing poverty, learning, and teaching through a critical race theory lens. *Review of Research in Education*, 37(1), 1–53. doi: 10.3102/0091732X12459720
- Moll, L. C., & Ruiz, R. (2002). The schooling of Latino children. In M. M. Suárez-Orozco & M. M. Páez (Eds.), *Latinos: Remaking America* (362-374). Berkeley, CA: University of California Press.
- Moschkovich, J. (2007). Bilingual mathematics learners: How views of language, bilingual learners and mathematical communication effect instruction. In N. S. Nasir & P. Cobb (Eds.), *Improving access to mathematics: Diversity and equity in the classroom* (89–104). New York, NY: Teachers College Press.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (1995). *Assessment standards for teaching mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Science Foundation. (1996). *Indicators of science and mathematics education 1995*. Arlington, VA: Author.
- Newmann, F. M., Bryk, A. S., & Nagaoka, J. (2001). *Authentic intellectual work and standardized tests: Conflict or coexistence*. Chicago, IL: Consortium on Chicago School Research.
- Nichols, S. L., & Berliner, D. C. (2008). Why has high-stakes testing so easily slipped into contemporary American life? *Phi Delta Kappan*, 89(9), 672–676.
- Nichols, S., Glass, G., & Berliner, D. (2006). High-stakes testing and student achievement: Does accountability pressure increase student learning? *Education Policy Analysis Archives*, 14(1). Retrieved from <http://epaa.asu.edu/ojs/article/view/72>
- No Child Left Behind Act of 2001. (2001). Retrieved at <http://www.ed.gov/policy/elsec/leg/esea02/index.html>
- Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd Ed.). New Haven, CT: Yale University Press.

- Olsen, B., & Sexton, D. (2009). Threat rigidity, school reform, and how teachers' view their work inside current education policy contexts. *American Educational Research Journal*, 46(1), 9–44.
- Payne, C. (2008). *So much reform, so little change: The persistence of failure in urban schools*. Cambridge, MA: Harvard Education Press.
- Payne, K. J., & Biddle, B. J. (1999). Poor school funding, child poverty, and mathematics achievement. *Educational Researcher*, 28(6), 4–13.
- Pegg, J. (2003). Assessment in mathematics: A developmental approach. In J. M. Royer (Ed.), *Mathematical cognition* (227–259). Greenwich, CT: Information Age Publishing.
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. Philadelphia, PA: Basic Books.
- Raymond, M. E., & Hanushek, E. A. (2003). High-stakes research. *Education Next*, 3(3), 48–55.
- Secada, W. G. (1992). Race, ethnicity, social class, language, and achievement in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (623–660). New York, NY: Macmillan.
- Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Suurtamm, C., Lawson, A., & Koch, M. (2008). The challenge of maintaining the integrity of reform mathematics in large-scale assessment. *Studies in Educational Evaluation*, 34(1), 31–43.
- Tate, W. F. (1995). Economics, equity, and the national mathematics assessment: Are we creating a national tollroad? In W.G. Secada, E. Fennema, & L.B. Adajian (Eds.), *New directions for equity in mathematics education* (191–208). New York: Cambridge University Press.
- Téllez, K., Moschkovich, J. N., & Civil, M. (Eds.). (2011). *Latinos and mathematics education: Research on learning and teaching in classrooms and communities*. Charlotte, NC: Information Age Publishing.
- Valenzuela, A. (Ed.). (2005). *Leaving children behind*. Albany, NY: State University of New York Press.
- Wilson, L. D., & Kenney, P. A. (2003). Classroom and large-scale assessment. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A Research Companion to principles and standards for school mathematics* (53–67). Reston, VA: National Council of Teachers of Mathematics.