HIGH-STAKES STANDARDIZED MATHEMATICS TESTING IN URBAN HIGH SCHOOLS: AN EQUITY PERSPECTIVE

Laurie H. Rubel
City University of New York (Brooklyn College)
LRubel@brooklyn.cuny.edu

This article examines mathematics opportunities at two urban high schools in the context of high-stakes standardized testing. Case studies were conducted at two urban high schools in low-income neighborhoods. Data includes opportunities in mathematics offered to students at the school level, in terms of mathematics course sequencing, and at the classroom level in terms of task cognitive demand and participation modality. In particular, this article focuses on two related questions: 1) Do high-stakes standardized testing set high standards for all students? and 2) Do mathematics teachers "teach to the test" and if so, how?

Educational equity is often framed as a significant rationale for high-stakes standardized testing (Diamond & Spillane, 2004). Proponents of high-stakes standardized testing argue that the tests hold schools and teachers accountable to providing all students with a high quality education (No Child Left Behind Act, 2008). Others (e.g., Au, 2007; Koretz, 2008) contend that an emphasis on testing prompts schools and teachers to adapt their curricula to "teach to the test." The primary goal of this paper is to explore the two sides of this argument, in the context of mathematics education at two urban high schools that serve Black and Latino/a students from low-income families. While other studies explore the impact of standardized testing on teachers' instructional practice by strictly using a methodology of surveying or interviewing teachers (e.g., Barksdale-Ladd & Thomas, 2000), this study examines the potential relationship between high-stakes standardized testing and equity in mathematics education empirically, with data from classroom observations.

The paper begins with a presentation of the research context and description of data sources. It continues with a presentation of results organized around the dual themes of standardized testing as promoting high standards for all students and teachers "teaching to the test." The discussion section summarizes the study's findings and poses questions for further research.

Methods

Research Context

Urban school districts, because of their size, typically serve a diversity of students, along racial, linguistic, religious, and socioeconomic dimensions. However, while a city may be diverse across many dimensions, it is typically an aggregate of smaller neighborhood units, which are often homogenous in terms of race and socio-economic class. African American and Latino/a students from the lowest income families tend to be clustered in schools in particular urban neighborhoods (Lipman, 2004). The research described in this paper is conducted in two such neighborhoods in New York City, each with about 130,000 residents: Bushwick is primarily Latino/a (split among Puerto Ricans, Dominicans, Mexicans, and Ecuadorans), and Brownsville, about a mile away, is primarily African American. These neighborhoods are among the ten lowest income neighborhoods citywide; more than half of families in Bushwick and more than 2/3 of families in Brownsville are in the bottom two quintiles of city income levels (Furman Center, 2008).

Harwood and Carver are pseudonymously named high schools located in these two neighborhoods. At Harwood, in 2009-2010, just over two-thirds of students self-identified as "Hispanic or Latino/a," and the other nearly one-third self-identified as "Black or African American." About 90% of the school's families were recipients of public assistance. Carver

Wiest, L. R., & Lamberg, T. (Eds.). (2011). Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Reno, NV: University of Nevada, Reno.

also serves students from low-income families; about 81% of its families were recognized as recipients of public assistance. In 2009-2010, about 90% of Carver's students self-identified as "Black or African American," and the remainder of students self-identified as "Hispanic or Latino/a" (statistics cited in this paragraph are from New York State Education Department, 2010).

Harwood and Carver were the two partner schools for the Centering the Teaching of Mathematics on Urban Youth (CTMUY) professional development project in 2009-2011. CTMUY is an NSF-funded, integrated research and professional development project, focused on improving classroom-level opportunities to learn mathematics provided to students at high schools in underserved urban neighborhoods (for more details about the professional development project, see Rubel, in press and Rubel & Chu, accepted for publication).

Data Sources

Data for this analysis includes school-level data, teacher-level data, and classroom-level data, described in more detail below.

School data. Ethnographic school visits were conducted at Harwood and Carver in 2008-2009 and 2009-2010 as part of the schools' participation in CTMUY. Interviews were conducted with assistant principals and principals, as part of the data collection for that larger project about each school's course offerings in mathematics and rationale for those offerings. In addition, student standardized testing data was gathered from annual state report card.

Teacher data. Three focal teachers at Carver and four focal teachers at Harwood completed a Likert-scale survey which included items, among others, about the degree to which they emphasize a variety of objectives, such as preparing for standardized tests, increasing students' interest in mathematics, or preparing students for further study in mathematics.

Classroom observations. As part of the teachers' participation in CTMUY, each of three focal teachers at Carver and four focal teachers at Harwood were visited for ten classroom observations across the 2009-2010 school year. Observations were conducted in four clustered rounds across the school year. Each teacher's set of ten observations was conducted in the same class period, with the same group of students. Fieldnotes were taken during the classroom observations and were then expanded into detailed narrative descriptions. Two teachers each had single observations that were highly atypical, either because of teacher illness or student attendance, and those two observations were dropped from the data set.

Each lesson's main mathematical task was identified and then classified in terms of its cognitive demand, using categories from Henningsen & Stein (1997). Cognitive demand can be low-level, if the task presented to students relies strictly on memorization or if it is a strictly procedural task, that does not offer connections to concepts, understanding or meaning. Tasks that have a high level of cognitive demand might be complex and non-algorithmic ("doing mathematics) or procedural in nature, however, the task's procedures connect explicitly to concepts, understanding or meaning.

A second quantitative measure of each observed lesson pertains to the various participation modalities offered to students. I utilized a set of categories adapted from Weiss, Pasley, Smith, Banilower & Heck (2003): listening; investigating or problem solving; discussing; reading, writing, or reflecting; using technology; or practicing skills. Each lesson was subdivided according to the various participation modalities offered to students, and those modailities were quantified in terms of relative minutes of instruction.

Results

Results are presented in terms of two themes. First, I explore the notion of high-stakes testing as a mechanism that promotes high standards for teachers and students, using the examples offered by Carver and Harwood. Next, I analyze project data to investigate if and how teachers at Carver and Harwood "teach to the test" in the context of high-stakes standardized testing.

High Standards and/or Gatekeeper?

In New York, to graduate from high school, students must pass a state standardized exam in mathematics that corresponds with state courses in Algebra, Geometry, or Algebra 2/Trigonometry. However, because of the actual, or perceived, nature of mathematics as a strictly cumulative discipline, in practice, the state's entry-level algebra exam plays the dual role of functioning as the 'high standard for all students' and as the gatekeeper of high school graduation. Table 1 contains the percentage passing rates on the state's entry-level algebra exam, for the classes of 2009 and 2010 at Carver and Harwood at the end of their four years of high school. As shown, at Harwood, in both the classes of 2008 and 2009, only about half of its students passed the entry-level algebra exam by the end of high school. As a result, Harwood did not meet the Adequate Yearly Progress (AYP) thresholds determined by the No Child Left Behind Act in either year. After two years of not meeting AYP, Harwood was then labeled a "school in need of improvement (SINI)" and was subject to a state quality review of their mathematics program. In 2010, their results improved significantly and exceeded the AYP threshold, with 72% of the graduating class passing the state's entry-level algebra exam.

Carver High School, on the other hand, had relatively stable frequency of passing the algebra exam: 65% of students in the class of 2008, 72% of students in the class of 2009, and 66% of students in the class of 2010 passed the algebra exam by the end of high school. Carver's results exceeded the AYP threshold in 2008 and 2009. Although a higher percentage of students in the class of 2010 passed a mathematics exam than the class of 2008, in 2010, because of the changing nature of the AYP thresholds, Carver did not make AYP in mathematics (all data is from New York State Education Department, 2008, 2009, 2010).

	Class of 2008	Class of 2009	Class of 2010
Carver	65% (98)	72% (93)	66% (92)
Harwood	48% (105)	52% (102)	72% (97)

Table 1. Percentage of students passing entry-level algebra exam

As we see in Table 1, the percentage of seniors at Carver who passed the state algebra exam remained roughly constant, with a slight increase in 2009 and a corresponding decrease in 2010. Harwood, on the other hand, showed dramatic improvement in 2010. One interpretation of these results is that, after two years of low performance on the mathematics exams, the sanctions imposed on Harwood pushed the school and teachers to focus their efforts on preparation for the algebra examination. In fact, at Harwood, in 2009-2010, the school structured their staffing resources so that the eleventh and twelfth graders who had not yet passed the algebra exam had mathematics class two periods each day. This evidence suggests that the mechanism of the state standardized test, and the sanctions that come with poor performance, pushed Harwood to focus its resources on its struggling students, by offering them double the amount of instructional minutes in mathematics.

A second interpretation of this data focuses on the mathematics opportunities that result from an institutional emphasis on a *minimal* requirement. In other words, the accountability system does not rate schools in terms of the learning opportunities they provide to all

students. Instead, the accountability system sanctions schools whose students do not meet a minimum requirement, in this case, passing an entry-level algebra exam. So at Harwood, in the 2008-2009 and 2009-2010 school years, students who did not pass the algebra exam at the end of their ninth grade year were placed in remedial "repeater classes" to retake the algebra course. Many students continued to cycle through the remedial classes all the way through their four years of high school. For instance, in the fall semester of 2009-2010, nearly 67% of all Harwood general education students were taking or re-taking the New York State entry-level Integrated Algebra course. So by the end of their four years of high school, while 72% of Harwood's students from the class of 2010 passed the state's algebra exam, many of these students (and of course, the 28% of the class who did not pass the algebra exam) were not given the opportunity to study geometry or any other mathematics as part of their high school education.

The case of Carver High School strengthens the interpretation of high stakes standardized testing functioning as gatekeeper. At Carver, during the 2008-2009 and 2009-2010 school years, students progressed through a sequence of Algebra, Geometry, Algebra 2/Trigonometry, irrespective of whether they passed the entry-level algebra exam. Those students who did not pass the exam were assigned to before school, after school, or Saturday sessions to continue to practice for that exam. However, once Carver did not make AYP in mathematics in 2010, Carver changed their course sequencing policy for their students in 2010-2011. While all students progressed from 9th grade algebra to 10th grade geometry, any eleventh or twelfth grade student who had not passed the entry-level algebra exam was removed from the mathematics course progression and tracked into a designated remedial "test-prep" mathematics class.

The analysis in this paper has, thus far, focused on school-level processes of course sequencing in the context of high-stakes standardized testing. In the next section, the analysis zooms in to the classroom level to examine the issue of if and how teachers at Carver and Harwood "teach to the test."

"Teaching to the Test"

One way to examine the potential relationship between high stakes standardized testing and teachers' pedagogical practices is to survey teachers. At the start of the CTMUY project, for example, all seven focal teachers at Carver and Harwood reported that they place moderate or heavy emphasis on preparing students for standardized tests. On the same survey, only three of the seven focal teachers indicated that they place moderate or heavy emphasis on preparing students for further studies in mathematics. Similarly, only three of the seven focal teachers indicated that they place moderate or heavy emphasis on increasing students' interest in mathematics. Even though this is an extremely small sample of teachers, these results suggest that these teachers take the accountability system of high-stakes standardized tests seriously, mirroring findings of Barksdale-Ladd & Thomas (2000). I expand upon these findings to illuminate how a teacher's emphasis on preparing students for standardized tests might be expressed in mathematics instruction. I build upon the definition of curriculum as presented by Au(2007) as including a) subject matter content knowledge; in this case, the mathematics content that is included in a curriculum, 2) structure or form of curricular knowledge; how mathematics is structured and presented within a curriculum, and 3) pedagogy; how that mathematical knowledge is communicated.

Mathematics content of curriculum. The high stakes standardized tests and the mathematics learning standards are both created and endorsed by the state and the school district, and they are therefore, assumedly, in direct correspondence. In the case of New York, high schools and their teachers are accountable to aligning the mathematical content of their 9th grade curriculum to the state's algebra standards. Proponents of high-stakes testing

argue that this standardizes the opportunities offered to students. So even in low-income neighborhoods like Bushwick and Brownsville, students are studying the same set of algebraic concepts and skills as everywhere else in the state. Seen in this way, "teaching to the test" pushes teachers to maintain high mathematics standards for their students by using the state's mathematics standards to guide the mathematical content of their curriculum.

Form of curriculum. In schools like Harwood and Carver, where students enter high school with a history of poor test scores, teachers are under enormous pressure for their students to achieve good results on high-stakes tests. Another interpretation of "teaching to the test," especially in contexts like these, is, therefore, that teachers will pattern the actual mathematical tasks they pose to students in the classroom, or the structure or form of mathematical knowledge, to precisely reflect the tasks that are offered on the standardized exams. By their very nature, standardized test tasks are limited to the types of questions that can be expressed in multiple-choice or short-answer form, have a single correct answer, and perhaps most importantly, can be completed in a very short time. These tasks typically require low-level forms of knowledge in that they tend to either require recalling a term or a executing a well-defined procedure (Williams, 2010).

I analyze this issue of "teaching to the test" first by examining the cognitive demand level of the mathematical tasks in the 68 observed classes at Carver and Harwood. Table 2 contains the distribution of cognitive demand level of these mathematical tasks in the 29 observed lessons at Carver and the 39 observed lessons at Harwood.

	Low-Level			High-Level	
	Memorization	Procedures without connections	Other	Procedures with connections	Doing mathematics
Carver	10	10		8	
Harwood	1	27	2	9	1
Total	11	37	2	17	1

Table 2. Task Cognitive Demand

As shown in Table 2, 50 of the 68 tasks were rated as low-level tasks, indicating that they were either memorization tasks (11), focused on procedures without connection to concepts, meaning or understanding (37), or were otherwise low-level (2). Only 18 of the 70 tasks were rated as high-level tasks, indicating primarily that they were tasks that focused on procedures with connection to concepts, meaning or understanding (17). In addition, 17 of these 18 high-level tasks were limited to only two of the seven focal teachers. In other words, the students in the classes of four focal teachers were offered only low-level tasks, in the 40 observed lessons of those teachers. These results strengthen the claim that teachers "teach to the test" by offering students low-level mathematical tasks, which correspond to the types of tasks found on standardized tests.

Pedagogy of curriculum. Another means of analysis as to if and how teachers are "teaching to the test" is to consider the forms of participation opportunities they make available to students in their mathematics classes. Table 3 contains the relative frequency of each participation modality offered to students in the 29 observed mathematics classes at Carver and the 39 observed mathematics classes at Harwood.

Wiest, L. R., & Lamberg, T. (Eds.). (2011). Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Reno, NV: University of Nevada, Reno.

Teacher-centered modes of participation, such as listening to the teacher, practicing skills and "housekeeping" (taking attendance, collecting homework, distributing textbooks) dominated the observed lessons at both Carver and Harwood, as shown in Table 3. In fact, in the observed classes of four of the seven focal teachers, these teacher-centered modes of participation occupied more than 80% of the instructional time (not shown in Table 3). Notably, teachers organized classes, on average, for students to participate by "practicing skills" 37% of the mathematics time at Carver and 32% of the time at Harwood. At both schools, the "practicing" consisted of teachers distributing worksheets with tasks modeled specifically on previous state test items. Teachers created these worksheets using a free, Internet-based application, which consists of a search engine that generates such worksheets in printable form. Student-centered modes of participation, like whole-class discussion, listening to other students, problem solving or investigating, using technology, or writing were offered much more infrequently in the observed lessons, on average: only 26% of the time at Carver and 42% of the mathematics time at Harwood.

Forms of Participation	Carver 29 classes 60 min each	Harwood 39 classes 48 min each
Teacher-centered		
Listening to teacher	29%	22%
Practicing	37%	33%
"Housekeeping"	8%	3%
Total	74%	58%
Student-centered		
Discussing	10.%	12%
Listening to students	2%	3%
Investigating	9%	22%
Writing	4%	4%
Using technology	1%	1%
Total	26%	42%

Table 3. Frequency of Participation Modalities

There is a relationship between the cognitive demand of a lesson's task and the participation modalities offered to students to complete that task. Low-level tasks, which typically require memorization of facts or execution of steps of a well-defined procedure, lend themselves to listening and practicing. The teachers in this study who exclusively offered low-level tasks to their students also nearly exclusively offered the teacher-centered forms of participation of listening to the teacher and practicing. In contrast, tasks with higher-level cognitive demands, in which students are challenged to detect patterns or make and communicate connections across mathematical representations, afford more varied opportunities for students. The teachers in this study who offered high-level tasks to their students also offered student-centered participation modalities like investigating or discussing.

Discussion

In this paper, I have presented dual themes related to mathematics education in a context of high-stakes standardized testing. In some ways, the current accountability system in New York forces high schools to maintain high standards for all students. While in the past, high school students could earn "local diplomas" without passing any state mathematics exams, current state mandates require that students pass a state mathematics exam in order to

graduate from high school. In the context of schools in low-income neighborhoods, this mandate has had an impact in that schools are now required to place their 9th grade students in algebra courses, as opposed to primarily placing students in remedial level courses in "pre-algebra" or "consumer mathematics."

Although most of the students at Carver and Harwood pass their schools' 9th grade mathematics courses, most of the 9th grade students do not pass the corresponding state exam. This presents the schools with a dilemma: should the students retake a course that they have already taken and passed, should the students proceed through the high school mathematics curriculum, or should the school restructure its entry-level course to stretch over multiple years? The accountability system does not reward schools for having students take more or more challenging mathematics courses, and instead, sanctions schools for having students not meet the minimum requirement. Therefore, I argue that the accountability system encourages schools to follow the first option, i.e. cycle students through remedial courses, with a test-preparation emphasis, until they pass the test. The effect that this has had at Carver and Harwood is that this entry-level mathematics examination is positioned as a "finish-line" towards graduation, instead of a starting point for continued success. Students need knowledge in geometry and algebra 2/trigonometry to succeed in college placement tests or in college courses, so this inversion of a minimum requirement to function as an end goal ultimately does students a great disservice.

Others have claimed that "the weight of the high stakes testing environment falls heaviest on the shoulders of low income students and students of color" (Au, 2009, p.3). More specifically, Jones, Jones & Hargrove (2003, p.115) posit that low income students "are hit doubly hard (by high-stakes standardized testing) – not only do they tend to have lower scores on high stakes tests that may block them from subsequent opportunities, but the instruction that they receive might actually be worse than the instruction that they received before the testing policy was implemented." This claim is substantiated, perhaps, by this data and analysis.

While it remains hypothetical, of course, as to what the mathematics instruction at Carver and Harwood might resemble without high-stakes standardized testing, the observed structure and pedagogy of the curriculum at both Carver and Harwood are in alignment with standardized testing in several ways. Focal teachers most often presented students with tasks of low-level cognitive demand whose form corresponds to the low-level form of the state exams, typically involving recalling a vocabulary term or short, well-defined procedure. In addition, these low-level tasks were presented in a classroom pedagogical environment that privileges teacher-centered forms of participation. Students were typically expected to participate in mathematics, throughout the school year, by listening (to the teacher) and practicing, using worksheets created from a bank of test items readily available to teachers on the Internet. Opportunities for students to investigate patterns, to solve non-routine problems, to use technology to discover relationships, or to write or reflect about a mathematical concept or process were far more limited in the observed classes.

This preliminary analysis suggests that the context of high-stakes standardized testing is complex and merits analysis with an equity lens. A common perspective is that high-stakes standardized testing promotes equity by setting and enforcing high standards in mathematics. This analysis has demonstrated that we also need to consider the ways in which high-stakes standardized testing "exacerbates inequities" (Diamond & Spillane, 2004). This study has demonstrated the particular ways that high-stakes standardized testing also functions as a gatekeeper in urban high schools. In addition, this study provides description and nuance to how teachers "teach to the test" by providing students with low-level mathematical tasks using teacher-centered forms of participation.

Wiest, L. R., & Lamberg, T. (Eds.). (2011). Proceedings of the 33rd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education. Reno, NV: University of Nevada, Reno.

Notes

This material is based upon work supported by the National Science Foundation under Grant Nos. 0742614 and 0119732. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Au, W. (2007). High stakes testing and curricular control: a qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267.

Barksdale-Ladd, M.A. & Thomas, K. (2000). What's at stake in high stakes testing? Teachers and parents speak out. *Journal of Teacher Education,* 51(5), 384-397.

Diamond, J.B. & Spillane, J.P. (2004). High stakes accountability in urban elementary schools: challenging or reproducing inequality? *Teachers College Record, Special Issue on Testing, Teaching, and Learning*, 106(6), 1140-1171.

Furman Center for Real Estate & Urban Policy. (2008). *State of New York City's Housing and Neighborhoods, 2008 Report*.

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28, 524–549.

Jones, M.G., Jones, B.D., & Hargrove T. (2003). *The unintended consequences of high stakes testing*. Lanham, Maryland: Rowman & Littlefield Publishers, Inc.

Koretz, D. (2008). *Measuring up: what educational testing really tells us*. Cambridge, MA: Harvard University Press.

Lipman, P. (2004). *High stakes education: inequality, globalization, and urban school reform*. New York: Routledge Farmer.

New York State Education Department. (2008, 2009, 2010). *School Report Card: accountability and overview report.*

No Child Left Behind Act of 2001, 20 U.S.C. § 6319 (2008).

Rubel, Laurie H. (in press, 2012). Centering the teaching of mathematics on urban youth: learning together about our students and their communities. In J. Bay-Williams and R. Speer (Eds). *Professional Collaborations in Mathematics Teaching and Learning: Seeking Success for All* (NCTM 2012 Yearbook). Reston, VA: NCTM.

Rubel, Laurie H. & Chu, Haiwen. (accepted for publication). Reinscribing urban: high school mathematics teaching in low-income communities of color. *Journal of Mathematics Teacher Education*.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E., & Heck, D. J. (2003). *Looking inside the classroom: a study of mathematics and science education in the United States.* Chapel Hill, NC: Horizon Research.

Williams, C. (2010). Algebra and No Child Left Behind: standardized tests and algebraic complexity. In Brosnan, P., Erchick, D. B., & Flevares, L. (Eds.). *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Columbus, OH: The Ohio State University (pp. 202-209).