EDUCATION POLICY ANALYSIS ARCHIVES

A peer-reviewed scholarly journal
Editor: Sherman Dorn
College of Education
University of South Florida

Volume 16 Number 3 January 31, 2008 ISSN 1068–2341

Avoidable Losses: High-Stakes Accountability and the Dropout Crisis[1]

Linda McSpadden McNeil
Eileen Coppola
Judy Radigan
Rice University

Julian Vasquez Heilig
University of Texas-Austin

Citation: McNeil, L. M., Coppola, E., Radigan, J., & Vasquez Heilig, J. (2008). Avoidable losses: High-stakes accountability and the dropout crisis. *Education Policy Analysis Archives, 16*(3). Retrieved [date] from http://epaa.asu.edu/epaa/v16n3/.

Abstract
In the state of Texas, whose standardized, high-stakes test-based accountability system became the model for the nation's most comprehensive federal education policy, more than 135,000 youth are lost from the state's high schools every year. Dropout rates are highest for African American and Latino youth, more than 60% for the students we followed. Findings from this study, which included analysis of the accountability policy in operation in high-poverty high schools in a major urban district, analysis of student-level data for more than 271,000 students in that district

[1] The authors gratefully acknowledge the contribution of Linda Darling-Hammond of Stanford University, who was the Principal Investigator on the part of the study funded by the Rockefeller Foundation. We also thank the Rice University Center for Education and Dr. Maconda Brown O'Connor for their support of this research. The authors are indebted to Laurie Hammons for her expert editing and analytical advice. The paper is stronger for the suggestions of Sherman Dorn, EPAA Editor, and the anonymous reviewers. The findings and analyses are those of the authors.

over a seven-year period under this policy, and extensive ethnographic analysis of life in schools under the policy, show that the state's high-stakes accountability system has a direct impact on the severity of the dropout problem. The study carries great significance for national education policy because its findings show that disaggregation of student scores by race does not lead to greater equity, but in fact puts our most vulnerable youth, the poor, the English language learners, and African American and Latino children, at risk of being pushed out of their schools so the school ratings can show "measurable improvement." High-stakes, test-based accountability leads not to equitable educational possibilities for youth, but to avoidable losses of these students from our schools.
Keywords: high-stakes testing; dropouts; accountability; Latino youth; African American youth.

Perdidas evitables: Acontabilidad escolar con exámenes de alto riesgo y la crisis del abandono escolar

Resumen
En el estado de Texas, cuyo sistema de acontabilidad escolar se basa en la utilización de exámenes de alto riesgo y que se convirtió en el modelo a seguir para la política educativa más exhaustiva del país, cada año más de 135,000 jóvenes se pierden al dejar de estudiar en las escuelas secundarias estatales. Las tasas de abandono escolares más altas se dan entre los estudiantes afroamericanos y latinos, que configuran más del 60% de los jóvenes investigados en este trabajo. Este estudio, incluyó el análisis del sistema de acontabilidad utilizado en un distrito escolar de gran magnitud, en sus áreas de alta pobreza, análisis de datos de más que 260,000 estudiantes durante un periodo de siete años, y análisis etnográficos extensos de la vida en las escuelas permitió obtener resultados que muestran una relación causal directa entre la severidad creciente de los exámenes y la problemática de los estudiantes que abandonan sus estudios. Este estudio tiene una gran importancia para la política educativa a nivel nacional porque sus conclusiones demuestran que la desagregación de las notas de los alumnos por raza no tuvo como resultado más equidad, sino que de hecho puso a los jóvenes de los sectores más vulnerables, pobres, estudiantes cuya primera lengua no es el ingles, afroamericanos y latinos, en mayor riesgo de ser expulsados de sus escuelas para que las mediciones de los rankings de las escuelas pudieran demostrar "aumentos apreciables". Sistemas de acontabilidad escolar basados en la utilización de exámenes de alto riesgo no lleva a mejorar la igualdad de posibilidades educativas , si no a problemas educativos para los jóvenes que podrían ser prevenidos.

Introduction

Each year Texas public high schools lose at least 135,000 youth prior to graduation. Despite the state's official dropout statistics, which hover in the 2.8–3.2% range (Texas Education Agency, 2002), 25% of Anglo students fail to graduate. The percentage of African Americans is closer to 45%, though greater than 50% for African American boys; and Latinos graduate at slightly under 50%. The graduation rates for the state's major urban districts, with the exception of El Paso, barely reach 50% of all youth (Balfanz & Legters, 2001; Haney, 2000; Intercultural Development Research

Association, 1999; Orfield, Losen, Wald, & Swanson, 2004). The discrepancy between the official rates and the actual disappearances can be partly accounted for in the state's method of counting: More than 20 *leaver codes* remove students from the grade cohort for such reasons as pregnancy, incarceration, declaring an intent to take the General Educational Development (GED) exam sometime in the future, or declaring an intention to transfer to another district. School officials can apply leaver codes for declared intentions even though there are no procedures to follow whether children do move into another educational setting (Losen, Orfield, & Balfanz, 2006). What is less easy to understand is how the problem of losing so many young people from the schools could persist after more than a decade of a comprehensive education reform policy whose purpose is to hold everyone in the education system accountable for producing measurable improvements. Furthermore, the state system includes a special provision for assuring that all children, including historically underserved ethnic minorities (in this state, Latino and African American), are assessed; their test scores are disaggregated and no school may be given a high rating if the scores of those subgroups do not show improvement. It was this feature of the Texas system, the disaggregation of test scores by race and the theory that paying attention to the patterns of the subgroups would narrow the achievement gap, that convinced other states and finally the President and members of Congress in both parties to believe that the Texas system of education reform should become the model for the nation. However, the large number of losses from Texas high schools well into the second decade of the state accountability system prompted the central question guiding this study: Is there a connection between high-stakes accountability and high numbers of dropouts?

The standardized, high-stakes, test-based accountability system in the state of Texas became the model for the nation's most comprehensive federal education policy, the Elementary and Secondary Education Act of 2002, commonly referred to as No Child Left Behind (NCLB). Considering the unmistakable relationship between Texas and NCLB accountability, studying Texas's first-generation accountability system provides an opportunity to probe the theory of action underlying NCLB: Schools and students held accountable to these measures will automatically increase educational output. The reality is far different. The findings of this study show that the accountability system itself is complicit in the very losses it claims to reverse. The losses are avoidable, but not while this accountability system governs schools.

High-stakes accountability

The Texas accountability system for public education is an extreme form of centralized management, with a strict hierarchy in which rules and sanctions are set at the top, with every level of the system accountable to the level above it for measurable performance. The most salient performance indicator, for all levels of the system, is the test score on the state's mandated standardized tests. The system was phased in over a number of years, first testing students in the 1980s in basic skills (Texas Assessment of Basic Skills, or TABS, for 1979–1983) and then a slightly higher level of skills (Texas Assessment of Minimum Skills, or TEAMS, beginning in 1984). The 9th grade class of 1987 was the first in which students were required to pass an exit level test (given first in 11th and later in 10th grade) to graduate. By 1990 the state had instituted the Texas Assessment of Academic Skills (TAAS), mandated to be given in key grades, with a primary focus on reading, math, and writing. (The key tested grade levels changed several times in the early years of the test.)

Through a decision of the Texas legislature, in 1993 that test became the linchpin of the accountability system, the first to have consequences for children, schools, and districts. In 1994, Houston Superintendent (and later U. S. Secretary of Education) Rod Paige made the test high stakes for school principals in the Houston Independent School District, basing principals'

performance evaluations on each school's test scores. Although other districts piloted variations in accountability incentives (most notably Dallas), the Houston district's model came to predominate. This change formalized the school as the primary unit of analysis in the accountability system. The claims for a centralized, standardized accountability system included improving achievement, narrowing the achievement gap between White and African American and Latino students, lowering the dropout rate, and increasing graduation rates for high school students. The single most important measure, the measure that was publicly reported, was the school-level pass rate on the state test. The test itself changed in 1999 (now the TAKS, Texas Assessment of Knowledge and Skills), and the graduation exit test has shifted to 11th grade. The use of the state test for administrative accountability remains in place, and science and social studies tests have been added in recent years.

The system is high stakes for administrators because they have been required to sign away their tenure, and they have no collectively bargained contract: Their potential for cash bonuses on the up side, or alternately for job loss, depends on the production of rising test scores at their school.[2] The principals' performance contracts are annual, thus producing great pressure to show test score increases within a very brief time period, far less time than actually needed to improve instruction, update school curricula, or enhance the capacities of teachers—in short, to make the kinds of investments that over time generate real improvements in instruction and learning. The stakes are high for students as well: Regardless of the number of credits and grades earned, students may not graduate from a Texas public high school without passing the state mandated exit test. The stakes recently became high for even very young children as well, with the addition of third- and fifth-grade testing requirements for grade promotion (again, without regard to the students' academic record) (Texas Code, 2007). The high stakes for student graduation and for administrators' contracts have been in place since the early 1990s. The extraordinarily high numbers of dropouts under such a system led us to investigate what happens to children in a standardized accountability system after it becomes normalized over a number of years. That normalization of standardized accountability shapes the way a state and its districts conduct school.

Proponents of this system have seen rising test scores and a narrowing achievement gap, leading them to dub the period under study the "Texas Miracle." It is true that official reporting of the accountability system's own indicators has consistently shown improvement in student performance on the state test, with rising scores, as well, for all three ethnic groups whose scores are disaggregated for purposes of rating schools: White, Latino, and African American. And as will be discussed below, there has been an apparent narrowing of the gap in scores between "White" and "non-White" subgroups. School scores have risen, with increases in the number of schools moving up the state rating system, from Acceptable to Recognized or even Exemplary (Texas Education Agency, 2003; Vasquez Heilig, 2006). That trend in the state's rating system contrasts with other indicators of achievement for Texas students (such as the SAT, ACT, the state college readiness test, NAEP), indicators not included in the accountability system and which declined or remained flat during this period (McNeil, 2005). Omitted from the school rating system are aspects of schooling essential to a quality education but not captured by the state test scores: adequacy of resources, quality of teaching, connection to families, and the breadth of sports and arts programs, among others. But those aspects were ignored in the Texas accountability system: Only the test scores on the state test mattered, and those were going up statewide. They were also going up in Brazos City, the city where this study was conducted and one of the first major cities in the state to make the

[2] There is no test of children's learning that has been designed to evaluate schools or school staff; the use of children's test scores for these purposes violates professional ethics; see American Educational Research Association, 2000.

system high-stakes for administrators by basing administrator contracts on student scores in their buildings.[3]

If the premise of the accountability system, its hierarchical controls, and the linking of measurable results to concrete rewards and sanctions were sound, then Brazos City should have been expected to lead the state in improved student achievement. Instead, year after year, national studies of urban districts showed Brazos City to be losing at least half its students before graduation (Balfanz & Legters, 2001, 2004; Haney, 2000; Orfield, et al., 2004; Young, 1999, 2002). More disturbingly, the disaggregation of test scores by racial subgroup did not seem to be leading to increased levels of graduation for these students. In fact, as will be seen in a later section of this paper, more seem to be leaving school and doing so at younger ages, many as early as 9th and 10th grade, generally too young for GED classes, work programs, or other constructive opportunities for continued education.

Based on years of close-up observation in Brazos City high schools, our early investigations revealed that the reliance on the school as the unit of analysis within the accountability system obscured the system's effects on children. Its focus on measurable outcomes rather than the quality of the education being provided to children had another effect: The obsessive focus on test scores also rendered invisible to the public the ways the system's incentive structures shifted school practice into the production of test scores, even when doing so reduced the quality of instruction (Coppola, 2007; McNeil, 2001; Radigan, 2007).

More was at stake than the principals' jobs. A reform so radical as a statewide, standardized system itself depended for its public legitimacy on continued evidence of improved results. Year after year, students in the lower grades would show greater gains than students in middle and high school, and students in lower grades also showed greater gains across the racial subgroups. At some point, the state's high school graduation exit exam would reveal weaknesses in the system that had not been addressed: inequalities of resources, inadequate support for English language learners (ELLs), and the test-drill system itself which had supplanted a more substantive curriculum in many schools.[4] Any observer could have expected that continuing inattention to resources would eventually hit the very visible wall of high school achievement: High school students had to pass a single exit test to graduate. Low scores in the high schools could jeopardize the legitimacy of the entire system.

The state's response was not a massive investment in the instructional capacity of the high schools. It was, instead, the creation of a legal loophole to permit principals to exclude from the tested cohort those students they deemed to be liabilities to the school-level scores, thus insulating the school's rating from the predicted scores of those students. This loophole was a waiver that permitted the alteration of grade-to-grade promotion rules. As explained in "Faking Equity: High-Stakes Testing and the Education of Latino Youth" (McNeil, 2005, p. 95), a school that the state approved for a waiver could hold back as a 9th grader any student who had failed even one semester of a core 9th grade course (English, math, science, social studies), regardless of the number of credits successfully passed. When administrators used this waiver to hold a child in 9th grade, retention excluded that student from the administration of the 10th grade TAAS test, the test on which the school ratings and the principal's job security depended. In some schools, the students had to re-take the entire 9th grade year, even courses they had passed; other schools assigned

---

[3] The names of the city, the school district, and the schools in this study are pseudonyms, as well as the names of individuals interviewed.

[4] In this paper we use ELLs to designate our immigrant students rather than Limited English Proficient (LEP). Limited English Proficient will only be used in the statistical section of this paper because it is a state-designated category.

students to "local credit" courses, including TAAS-prep courses, where students earned no credit toward grade-level promotion or graduation. For others who repeated 9th grade because administrators applied this waiver, students were allowed to take 10th grade courses in other subjects but were officially on the rolls as 9th graders, excluded from the tested cohort until they passed the course they had failed. Not all schools placed retained 9th graders immediately in the courses they had failed.

As will be discussed in the case study analysis of Edgeview High School in a later section of this paper, the waiver provided cover for principals, whose performance was judged on short-term test score results; but it did not exempt students from eventually having to pass the high-stakes test to graduate. Interviews with Brazos City principals and teachers brought to light their awareness that as many as one-third to one-half of their students might be retained under the waiver, greatly inflating the subsequent year's test scores. The predictable fact that many or most of the retained students would leave school before graduating or even attaining 10th grade status to be test-eligible was known at the time and still allowed.

This waiver was the first example of what we have come to understand through this study as systemic incentives to take administrative action that encourages low-scoring students to leave the system, to drop out of school. We encountered this waiver as a mechanism for student exclusion as a dilemma it posed for principals, not as a matter of policy. Principals told us that taking advantage of this loophole could raise their school ratings: It enabled them to hold their weakest students back a grade, keeping them out of the tested cohort, out of jeopardizing the principal's performance contract. They also knew that holding students back a grade level is almost assuredly a ticket to dropping out, especially if a student is only one-half credit short of attaining the next grade level (Jimerson, Pletcher, Graydon, Schnurr, Nickerson, & Kundert, 2006).

Along with other stringent provisions of the accountability system, such as the criminalization of minor student offenses and attendance policies that prevented earning course credit after only a few absences, this waiver provision made it clear that the public face of the accountability system—the school as the unit of analysis—was inadequate for understanding the way the system works in schools and the ways it impacts children. If we were to understand the impact of the system on children, including the possibility that it was leading to more dropouts, we would need to shift the unit of analysis to the student. The invisibility of the losses of thousands of youth from the schools, because they fall outside the accountability system's indicators, was reason enough to expand our investigation of the effects of the system and to shift the analysis to the children, to their experience under this system.

From the system to student data and back into schools

Texas was the first state to implement high-stakes, test-based accountability throughout its system. It is the second most populous state, and its diverse population increasingly presages the population of the nation's school children. With its reliance on testing children in order to assess schools and school staff, the Texas accountability system was the basis for the Elementary and Secondary Education Act reauthorization in 2002 (as NCLB). Now that the Texas system has been in place for almost 15 years, it offers a compelling case study tracing the effects of this system, its rewards and sanctions, its definitions of improvement and achievement, and its view of children after such a system has been operationalized and normalized over time.

A question that kept arising in the schools was the possible connection between dropping out and the accountability system's pressures to produce high school-level scores. We performed a statistical analysis to compare performance under the accountability system's own indicators, which

employ school-level scores, to student progress through school using student-level data. This is presented below in "Uncovering the Connection between High-Stakes Testing and the Loss of Minority Students." Principals seemed to take this connection for granted and were puzzled that we thought it ought to be investigated. It was at the heart of their dilemma *to comply or educate*. In a later section of this paper, "A Case Study in Compliance," we investigate this connection. Our analysis of student-level data, over a period of years in which the accountability system was fully in place, showed a clear relationship between the disappearance of increasing numbers of students and schools' rising accountability ratings.

To investigate how and whether the accountability system's own incentive structures might be contributing directly to students' dropping out, or whether the correlation arose from cheating or other factors, the research team followed one school over an extended period of time, documenting the decisions made by its administration and faculty, including the invoking of the 9th grade waiver, and their impact on students in this high-poverty, mostly Latino high school. The findings within this school, confirmed in interviews with staff at the other schools whose patterns of advanced rankings similarly parallel greater 9th grade retention (students being held back in that grade) and increased dropouts, showed that the connection between the system's demands and dropping out is not merely coincidental. These schools are complying with the system and earning its rewards at the same time they are losing students. This school case study documents the processes and the logic by which this occurs. The testing component is only the most visible and publicly discussed aspect of the accountability system. We document other high-stakes practices under this system which individually, collectively, and cumulatively work against progress through school for many students.

Emerging Research on Standardized Accountability

The need for a systemic, longitudinal study that traces out the impact of accountability policies to the students would be justifiable in light of the magnitude of the losses of youth from Texas schools alone, but the problem is not limited to Texas. Several previously published studies have found higher rates of retention and dropping out in states and cities that have instituted tougher graduation requirements, even as test scores improve (Lillard & DeCicca, 2001; Clarke, Haney, & Madaus, 2000; Orfield & Ashkinaze, 1991; Roderick & Camburn, 1999; Wheelock, 2003). An earlier study analyzing the National Educational Longitudinal Survey (NELS) data from 1988 showed little association into the early 1990s between mandated high school exit exams and an increase in the number of students taking the GED or dropping out of high school (Warren & Edwards, 2005). Our findings of a strong connection are from a time period in which the exit tests had become institutionalized, and the accountability connection to administrative incentives had been operationalized. Using individual-level data from the NELS, Jacob (2001) found that graduation tests increased the probability of dropping out among the lowest-ability students. These responses to testing contribute to a widening gap in graduation rates between White and minority students (Orfield et al., 2004).

Smith (1986) explained the widespread engineering of minority student populations that he found in his study of New York City's implementation of test-based accountability as a basis for school-level sanctions:

> (S)tudent selection provides the greatest leverage in the short-term accountability game… The easiest way to improve one's chances of winning is (1) to add some highly likely students and (2) to drop some unlikely students, while simply hanging on to those in the middle. (pp. 30–31)

More recent evidence suggests that many of New York City's high schools may be trying to improve their test scores by pushing out weaker students who are unlikely to pass the state's high school graduation tests, first imposed in 1999. By 2000–01, more than 55,000 high school students were discharged without graduating, a number far larger than the 34,000 seniors who graduated from high school (Gotbaum, 2002), and the number of school-age students in GED programs run by the city schools increased by more than 50% as the tests were phased in, from 25,500 to more than 37,000 (Arenson, 2004). Fine (2005) notes specifically that fewer than 40% of the Black and Latino students were graduating in 2005. She adds that schools are removing students from their records, "so much so that Franklin Lane High School was sued for systematic removal of 'difficult children' (50% or 1,600 of its students each year)" (p. 25).

In their study of Chicago Public School sixth- and eighth-graders who had been retained, Jacob and Lefgren (2007) found that being retained in sixth grade made little difference in whether these students dropped out or the ages at which they dropped out, when compared to their peers who had been promoted. By contrast, students retained in 8th grade were 14% more likely to drop out. This suggests that grade retention late in a student's academic career can have a significant impact on academic attainment. Their study is directly related to high-stakes accountability in Texas because of two parallel mechanisms. First, the criteria for grade promotion were based on passing the high-stakes reading and math tests, rather than a comprehensive review of the student's academic record. Two, being retained at a later grade may create social dislocation, as well as give students less time to catch up with their peers academically.

In Massachusetts, which began requiring a high school exit exam for graduation in 2002, graduation rates decreased from 76 percent in 2002 to 72 percent in 2003. Meanwhile some of the steepest increases in test scores occurred in schools with the highest grade-retention and dropout rates. For example, high schools receiving state awards for gains in 10th grade pass rates on the Massachusetts test showed substantial increases in prior-year 9th grade retention rates and in the percentage of "missing" tenth graders (Wheelock, 2003).

Orfield and Ashkinaze's (1991) study of Georgia's testing policies coupled with Atlanta's 1980s test-based pupil promotion policies found that, as predicted by previous research on the relationship between grade retention and dropping out, school completion rates dropped from over 70% to 65% by 1982 and 61% by 1988. Studies in the collected volume, *Raising Standards or Raising Barriers? Inequality and High-Stakes Testing in Public Education* (Orfield & Kornhaber, 2001) document an emerging pattern of negative consequences for educational equity accompanying the implementation of high-stakes, test-based accountability, a pattern across a number of states. The emergence in other states of this pattern of simultaneously rising test scores and dropout rates pointed to the need to investigate at the systemic level a possible connection. In his Texas study, for example, Haney (2000) suggested that one of the primary mechanisms by which students were exempted from testing in Texas-style accountability reforms of the 1990s was the use of the 9th grade as a holding bin prior to the 10th grade high-stakes testing year, with many students never progressing past the 9th grade year. Carnoy and Loeb (2002), on the other hand, argue that Texas's graduation rates did not decline as a result of high school exit exams. This study will help to resolve the conflict between these two studies with its quantitative and qualitative longitudinal analysis.

As Dorn argued in "High-Stakes Testing and the History of Graduation" (2003), the introduction of a high-stakes graduation exit exam does not clarify the value of the diploma as a proxy for students' learning, for school improvement, or for certifying the learning of specific skills (nor the quality of the educational program itself). He found that the historical ambiguities associated with a high school diploma remain under a high-stakes testing system. The social meanings of a diploma are complex and need to be disentangled before students, teenagers, are made to bear the potentially negative consequences of a high-stakes barrier to receiving it. As Dorn

(2007) has more recently claimed, the "blunt systemic behavior" of high-stakes accountability functions "as a cudgel, not a scalpel" with this complexity (p.76).

Methods

This study is able to capture the links between high-stakes, test-based accountability and student dropouts by stepping outside the assumptions central to much research on these topics. In each phase of the study, the central question was not how and whether the accountability policy is working, nor how dropout formulas could be more finely tuned, but what is happening to the youth themselves under a centralized, standardized accountability system. In "Official and Unofficial Stories: Getting at the Impact of Policy on Educational Practice," McNeil and Coppola assert that "because educational practice is highly complex…the analysis of the impact of policy on practice must itself be complex, capable of drawing on more than one theoretical framework, employing multiple methodologies, analyzing multiple levels of governance and organization, and building a rich and varied body of data" (McNeil & Coppola, 2006, p. 681). Good policy analysis, the authors argue, "needs to encompass not only the official versions and intended impacts, but also unofficial versions that take into account such factors as varied definitions of what constitutes accurate data, unequal power relations, and cultural interpretations." Such an analysis necessarily steps back from the official categories and assumptions of the policy to problematize them, to analyze their origins and embodied intentions. And such an analysis investigates the policy in operation, the policy as it is lived by the adults and children within the schools under its governance. Other writers as well have argued that complementary methods allow us to examine a phenomenon and its causes (e.g., Johnson & Onwuegbuzie, 2004; Maxwell, 1996; Miles & Huberman, 1994).

Data sources from macro to micro

The study reported here satisfies these methodological requirements. To triangulate the results, the study was conducted in four phases: an ethnographic study of urban high schools in Brazos City, an early adopter of high-stakes accountability; a statistical analysis of a large, urban, Texas district's student-level data over a seven-year period; an in-depth ethnography of a high-poverty mostly Latino high school working to balance accountability compliance and educational improvement; and school-site interviews and observations in a larger sample of Brazos City high schools to investigate school-level actions along points of significance emerging in the statistical data. Because observational school-site data asserted direct links between the accountability system's reward structure and dropouts, we sought to expand the qualitative sample size and utilize complementary methodologies to obtain more generalizable results.

The Macro. The quantitative portion of the study utilizes a longitudinal dataset of 271,000 students to examine the effects of the first-generation Texas-style accountability system in a large urban district from 1995–2002. This dataset covered a seven-year period with school and individual results from various tests, including the state's mandated assessment test. With this information, the descriptive analyses track individual grade-to-grade progression and whether students graduated. The inferential statistical analysis utilized a school-level data set aggregated from BCSD's individual-level data set. The school-level dataset contains data for twenty-four traditional high schools, after dropping a number of high schools from the analysis for lack of pertinent data—either from their recent establishment in the district or their unique nature as alternative schools for special populations (for example, a high school for expectant mothers). We analyze accountability rating

trends for the population of twenty-four traditional high schools arranged in a panel format where each individual high school may contribute to several records. As a result, schools and years are the units of analysis, as is consistent with the states' use of the school as the units of analysis in the accountability system.

To consider trends in BCSD high schools' accountability ratings, this paper utilizes multinomial logistic regression (see Table 3 in a later section), which estimates the probability of a specific event occurring and allows consideration of more than two categorical dependent variables. Using the same predictor variables, regression coefficients were obtained for three contrasting situations: a decrease in TEA school rating (used as the reference group), no change in rating, or an increase in the school rating. TEA ratings were a function of increases in Texas Learning Index (TLI) scores coupled with officially reported dropout rates below threshold levels for each rating. The model is as follows:

$$\log (\pi_j/\pi_J) = \alpha_j + \sum^{K}_{k=1} \beta_{jk} X_k$$

Independent variables are denoted by $X_{k;}$ these influence the probability π_j that category *j* of the response variable will be chosen. In this analysis, "1" is used as the reference category which denotes schools that remained at the same accountability rating. This analysis tests the relationship between year-to-year school-level changes in TEA accountability ratings and changes in student progression, demographics, and teacher capacity.

The independent variables control for changes in school-level demographic variables and measures of teaching capacity, including year-to-year changes in student characteristics (% White students, % LEP students, % special education students, % at-risk students) and teacher characteristics (% of teachers certified, % teachers with less than three years of experience, % annual teacher turnover). Each year-to-year change represents a separate observation in the multinomial regression models. (See Appendix A for descriptive statistics for variables used in the analysis.) Year-to-year change variables for student progress, school capacity, and student demographics were calculated as follows:

$$\Delta V_t = V_t - V_{t-1}$$

The Micro. Qualitative data was gathered through interviews and focus groups conducted with a total of 122 students, as well as 38 administrators and teachers in seven different schools. The schools sampled for the qualitative work were chosen to represent several demographic trends in the district. Overall, the schools in BCSD tend to be high poverty, and either majority African American or majority Latino, with several more mixed along racial lines. Therefore, we chose three majority African American schools, three majority Latino schools and one more mixed-race, all with relatively high proportions of students receiving free or reduced-price lunch. The chart below shows the pseudonyms and average demographics grouped within each type of school. To preserve confidentiality, we do not show individual school demographics. All schools and persons interviewed were assigned pseudonyms for the purpose of reporting this research in accordance with standard practices for maintaining the confidentiality of research subjects.

Table 1
Sample of schools for qualitative data collection

Schools	African American	Asian	Latino	White	LEP	Special Ed.	At-Risk	Free Lunch
Majority African American:	87%	1%	12%	1%	3%	20%	60%	76%
Douglass								
Carver								
Lincoln								
Majority Latino:	13%	2%	80%	4%	25%	13%	81%	88%
Edgeview								
Clearbend								
Crockett								
Mixed:	55%	1%	43%	1%	13%	18%	75%	71%
King								

The effects of the accountability system were explored through parallel interview protocols with administrators, teachers, regular-age students, and students who had dropped out and then returned to school. In these protocols, the researchers explored themes raised by the quantitative analysis for the period of 1996–2002 to discover the mechanisms by which rising test scores and accountability ratings appeared beside higher grade retention and dropout rates. We asked students about their experiences with high-stakes testing, with grade retention, and the curriculum they experienced. With former dropouts, we explored the reasons and stories behind their dropping out and how they decided to return to school. With administrators and teachers, we explored how they experienced high-stakes testing and accountability processes, whether and how the system affected the curriculum as they taught it, and how they perceived the effect of this cluster of policies on students. All interviews and focus groups were approximately one hour long, taped and transcribed, and then coded for consistent themes (Miles & Huberman, 1994; Patton, 1990). Further, we conducted observations of each school as part of data collection. One researcher from our team had observed periodically at Edgeview for approximately ten years; another member for more than seven years. A different member had been observing Clearbend for more than four years. At the remaining five schools, teams of two to four researchers visited the school for two to three days to conduct interviews and focus groups, while at the same time, we observed in offices, hallways, teachers' lounges, libraries, cafeterias, and other public spaces. At the five schools, some classroom observations were conducted as the opportunity arose. Field notes were taken at each visit and the observed data discussed extensively among the researchers to identify strong, consistent themes.

Policy mechanisms at the school level

Our ethnographic investigation leading up to the present study had suggested that retention of large numbers of 9th graders (up to one-half of the class in some schools) had its basis in the accountability system's incentive structures for principals, particularly their need to show annual increases in school-level test scores. The research team returned to schools to investigate the link between 9th grade retention, dropping out, and the incongruously rising school ratings for schools losing the greatest numbers of students. The investigation was deepened to provide a comprehensive, multi-year analysis of the ways the policy played out in a particular urban high school that fit the pattern of rising ratings and rising dropouts. The explanatory power of such a

close-up examination of organizational processes is rooted in our tracking practices and determining which were traceable to the accountability system.

The study was widened to include a larger sample of seven high schools, through our interviews with teachers, administrators, and students. This phase also included a re-examination of the accountability system's internal mechanisms, including not only the testing and school ratings, but incentive structures for administrators and components which were less visible, but experienced by the youth themselves often as punitive and, in many cases, barriers to persistence in schools. This phase of the study, then, connected the dots, showing that 9th grade retention and large numbers of dropouts are not anomalies in schools with rising ratings but instrumental in the production of those ratings. The interaction between intensive school-level ethnography and statistical analysis of large-scale student-level data, reinforced by extensive interviewing, showed how school staffs and students made decisions on the basis of the accountability system's requirements.

Uncovering the Connection between High-Stakes Testing and the Loss of Minority Students

To investigate what is happening to the youth themselves under a centralized, standardized accountability system this section of the paper utilizes a statistical analysis of a unique longitudinal dataset of 271,000 students to examine the effects of the first-generation Texas-style accountability, which operated from 1995 through 2002. The major focus of this research is to evaluate empirically whether such policies actually triggered changes in student achievement, or increased grade retention, dropouts, and disappearance of students from school in an urban district setting.

The Texas miracle in Brazos City

Achievement gains across grade levels, conjoined with increases in high school graduation rates and decreases in dropout rates, brought nationwide acclaim to the Texas accountability system. As a result, the Texas system of high-stakes testing and accountability rating formulas became a model for nationwide high-stakes testing and accountability policy.

Accountability ratings. During the formative years of the Texas accountability system, the primary base indicators for determining a high school's accountability ratings were the Exit TAAS scores and the annual dropout rate. Based on these measures, the following four Texas Education Agency (TEA) Accountability rating levels were assigned: Exemplary, Recognized, Acceptable, or Low-Performing. Although 26% of schools were rated low-performing in 1996–97, by the next year, no schools were rated low-performing. From 1997–98 to 2000–01, the proportion of schools falling in the top two categories (recognized and exemplary) increased from 8% to 43%. Thus, the myth of the Texas Miracle was born. After 2001, a shift in ratings occurred as TAAS school-level requirements for TEA Accountability ratings became more stringent. When the mandatory passing rates increased from 30 to 50% for a school to be rated Acceptable and from 75% to 80% for a rating of Recognized, suddenly most schools were rated low-performing once again.

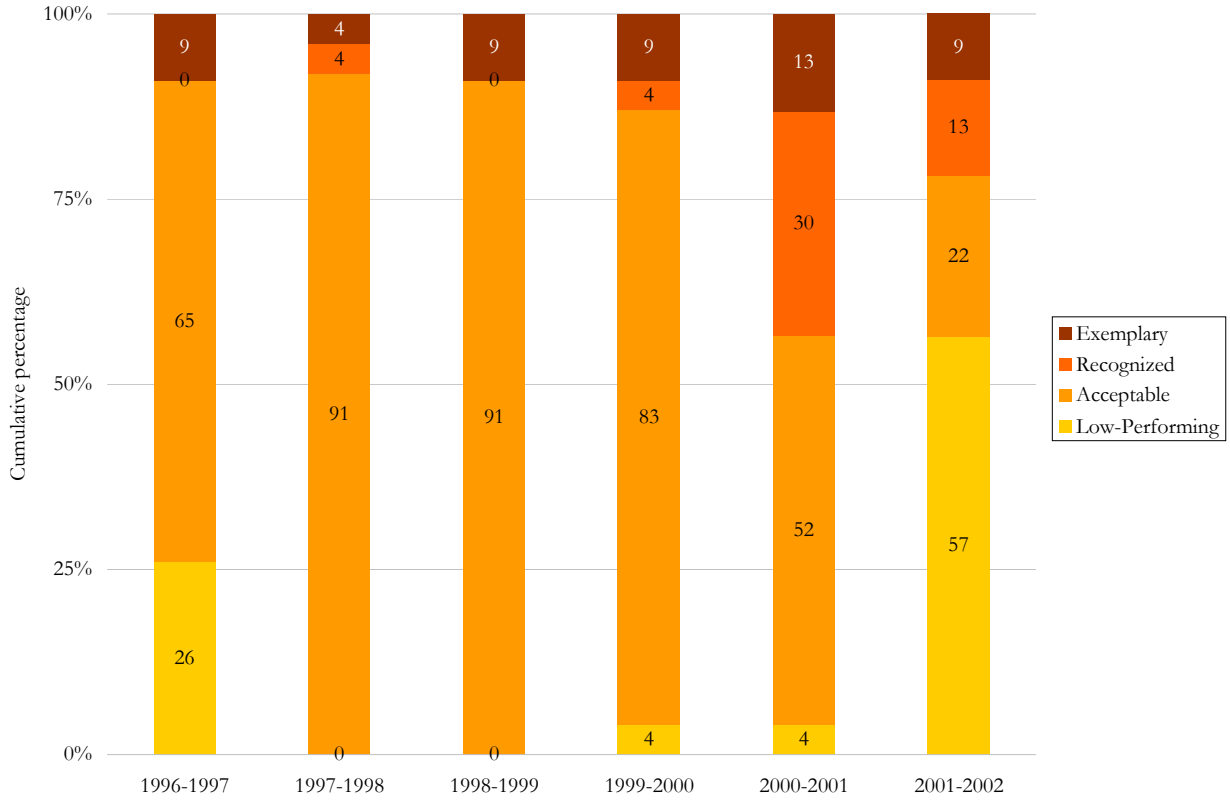


Figure 1. Percentage of BCSD high schools in each TEA accountability rating, 1996–2002.

The years from 1997 to 2001 show a significant shift in the number of BCSD high schools from Low-Performing to the Acceptable label. Many schools migrated to TEA Recognized and Exemplary status by 2000–01 (30% and 13%, respectively). Without data on the individual student trends, one might think the overall TEA Accountability rating trends show that traditional BCSD high schools extraordinarily increased their students' academic and enrollment success.

Dropouts. Haney (2000) described the decrease of the official rate of students dropping out of school as more evidence contributing to the perception of dramatic educational gains in Texas. The fact that a great many students fail to complete high school has plagued our educational system for decades, so the apparent success of accountability to stem the tide of dropouts statewide and, more specifically, in Texas' large urban districts showed promise. Figure 2 shows dropouts as officially reported by BCSD schools to TEA by grade and year.

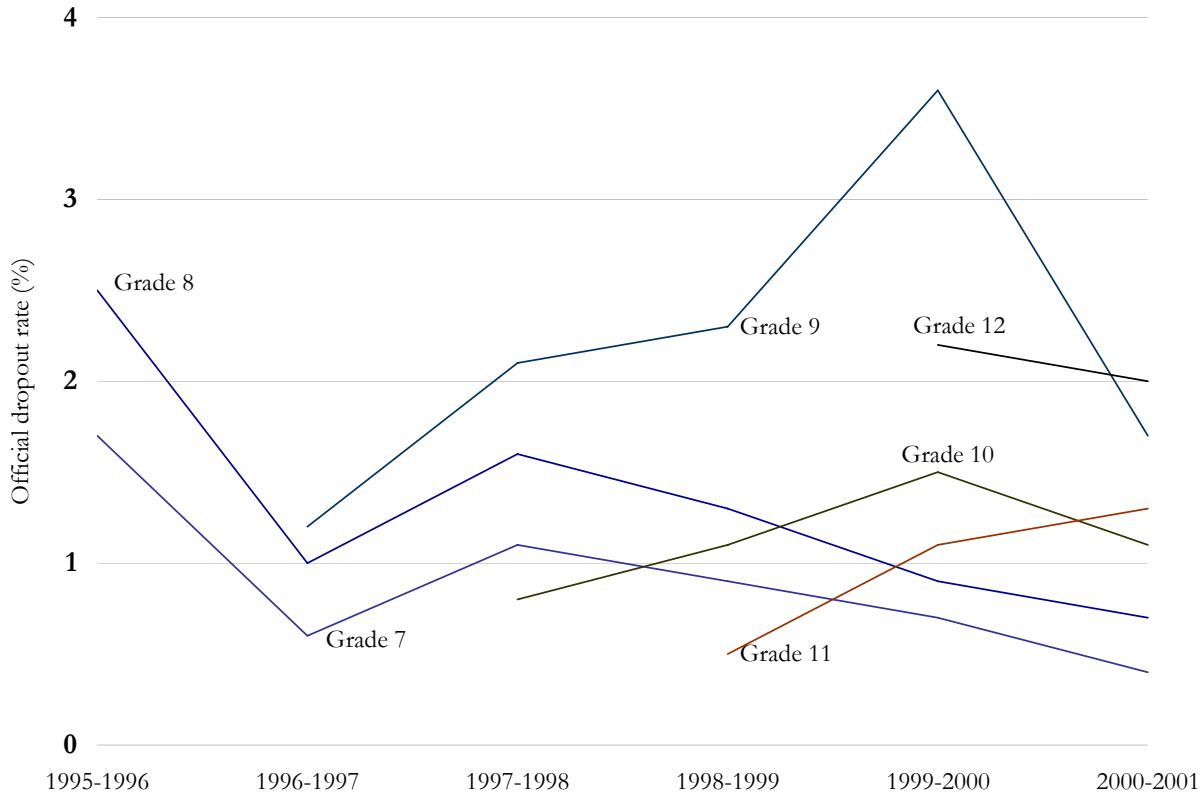


Figure 2. BCSD official dropout rates by grade, 1995–96 through 2000–01.

School-reported dropout rates in BCSD hovered below 2.5% for all grades except for the 9th grade, which peaked around 3.5%. Grades 10–12 are phased into the data set after 1996–97. Except for 11th and 12th grades, all grades show overall decreases in dropouts. This trend is also observed in Figure 3 below, which disaggregates official dropout rates by racial groups.

Figure 3. BCSD official dropout rates by race/ethnicity, 1995–96 through 2000–01. Reported for grades 7–12.

Figure 3 shows that school-reported dropout rates for grades 7–12 were below 2.5% for all ethnicities. Dropout rates for African Americans and Latino students were higher than Asian and White students. Similar to the grade analysis, dropout rates decrease over time for every ethnicity.

Graduation. Ultimately, the most important student outcome measure for high schools is the graduation rate. Concern about accountability ratings, dropout rates, and test scores are ancillary (though integral) to increasing the proportions of student receiving high school diplomas. Haney (2000) cites rising official graduation rates as an important measurement of the "miracle" educational gains in Texas. The Texas Education Agency report, "Secondary School Completion for 2000–01," claims that the state had an overall graduation rate of 81.1%. BCSD reported in 2002 that the district-wide graduation rate had climbed 21 percentage points over five years, from 54% in 1997 to 75% in 2002.

The district-level evidence we have considered thus far suggests that the system of accountability and high-stakes testing held schools in BSCD accountable and led to increased students' educational output as graduation rates climbed. In the public eye, dropout rates plummeted and test scores soared. The officially reported data appears to show that educators and students facing sanctions and rewards tied to batteries of high-stakes testing and accountability formulas worked harder and more effectively to enhance student learning and success.

The Texas miracle in Brazos City revisited

The factors that created the illusion of a Texas Miracle, of progress on overall achievement and of narrowing the achievement gap, are the results of the avoidable loss of a great many students, mostly poor and minority, from the testing pool. To uncover the enormity of the loss of these students, we followed the 1997 9th grade cohort containing over 13,000 students who were in the 8th grade in BCSD for 1996–97 and then the 9th grade in the following year.

Ninth-grade retention. The sudden growth in 9th grade retention rate from 1998 to 2001 for minority students is reflected in Figure 4 below. Carnoy, Loeb, and Smith's (2001) claim that that Texas's retention problems predate the TAAS testing is challenged, as McNeil's (2005) argument that ties accountability policy to increases in grade retention proves accurate. To make clear the longitudinal retention trends of students of color in BCSD, Vasquez-Heilig disaggregated BCSD retention data by Race/Ethnicity (See Figure 4). The data allowed for the examination of five years of 9th grade retention and reflected the growth of the minority student retention rate from 25% in 1997 to almost 35% in 2001.

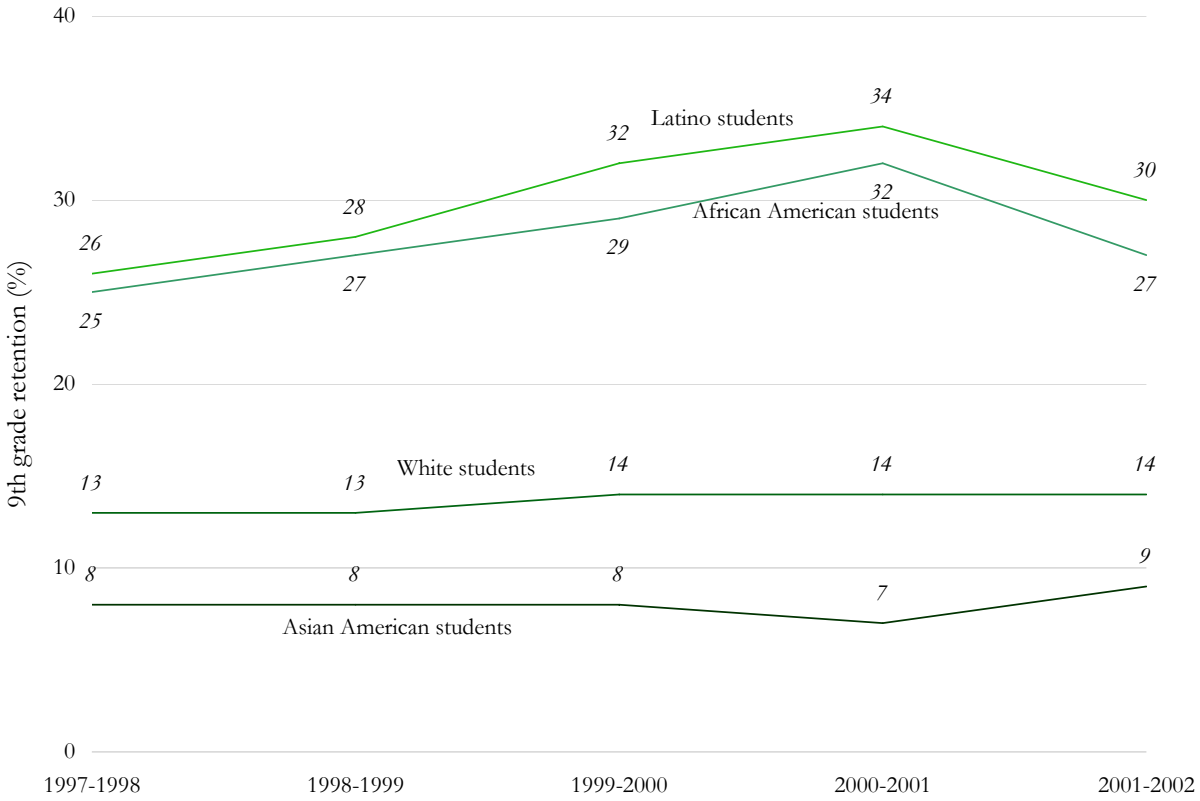
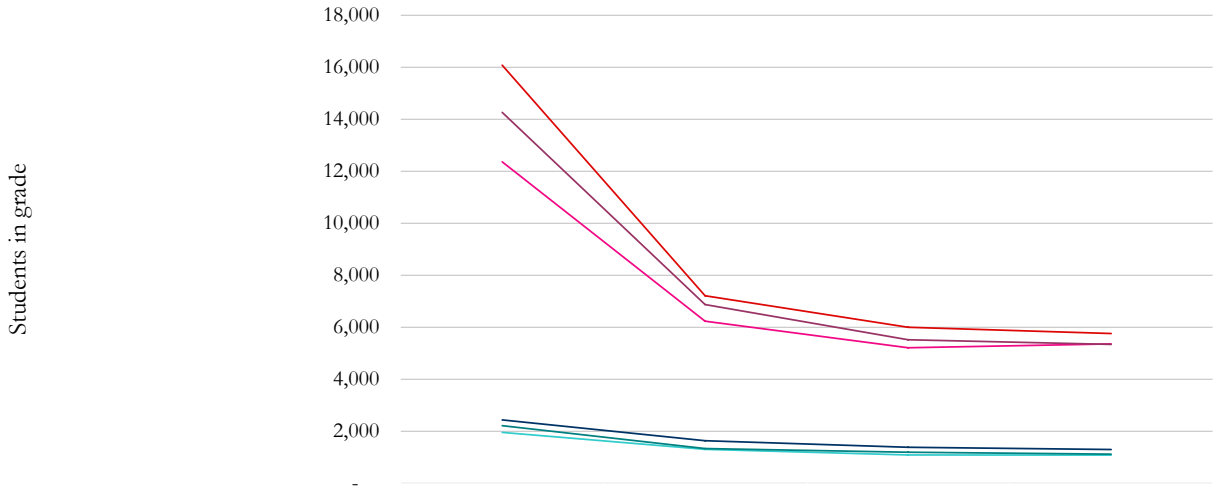


Figure 4. BCSD 9th grade retention by race/ethnicity, 1997–98 through 2001–02.

The figure shows that at the same time there is a claim of a narrowing of the achievement gap in TAAS testing for minority students, there is a large gap in 9th grade retention by Race/Ethnicity. African Americans and Latinos had the highest rates of retention as more than a quarter of these students were retained. Additionally, African Americans and Latinos also experienced increasing rates of retention to almost 35% for most of the period. In contrast, Asians and Whites had relatively flat rates of retention during the same period. What caused this sudden

peak in 9th grade retentions for at least three school years? A promotion standards (waiver) policy required freshman students to have a firm grasp of the basics and pass four core courses (algebra, English, science, and social studies) before being promoted to 10th grade. We entered schools to confirm the impact of the waiver on 9th grade retention and the concomitant false narrowing of the achievement gap and rise in high-stakes testing scores. In a following section, we provide a representative case study of how the accountability system affected students' grade progression, persistence to graduation, and the school's attempts at reform during this period.

High school grade progression. What effect did this grade retention have on students' progression through high school? Through the BCSD data set, we followed three entering 9th grade cohorts over a four-year period to see the effects of grade retention on high school progression. Figure 5 shows a yearly attrition of 6,000–9,000 African American and Latino students who did not progress on time from the 9th to the 10th grade. African Americans and Latinos show the steepest loss between the 9th and 10th grades as about half of the freshman class did not progress on time.



	9th grade	10th grade	11th grade	12th grade
1996 African American/Latino	12,361	6,229	5,212	5,363
1997 African American/Latino	16,076	7,212	5,997	5,757
1998 African American/Latino	14,265	6,877	5,517	5,344
1996 White/Asian	1,959	1,303	1,095	1,084
1997 White/Asian	2,216	1,340	1,196	1,118
1998 White/Asian	2,442	1,640	1,384	1,301

Figure 5. BCSD high school cohort progression by race/ethnicity (raw counts). Lines correspond to separate entering ninth-grade cohorts from 1996 to 1998.

The lack of grade progression for minority students is in contrast to the progression trends of White and Asian students. Only 600–900 White and Asian American students did not progress to the 10th grade on time. This translates to more than 60% of Whites and Asian American students arriving in 10th grade on time (see Figure 6 below). The graduation rate of students and the avoidable loss of students become more apparent as we study the cohorts' progression through four years of schooling. Of the three combined minority cohort groups, at most 40% were enrolled in 12th grade on time, while in the three White and Asian cohorts, more than half of students were

enrolled in 12th grade on time. Of note, these graphical analyses track only whether high students in BCSD arrived in grade on time during the typical four-year time frame. They do not track longer time periods or detail various methods of attrition such as transfer, death, or official school withdrawal and dropout. Vasquez Heilig and Darling-Hammond (in press) provide an extended cohort analysis examining high school students who officially withdrew, dropped out or otherwise left BCSD high schools over time.

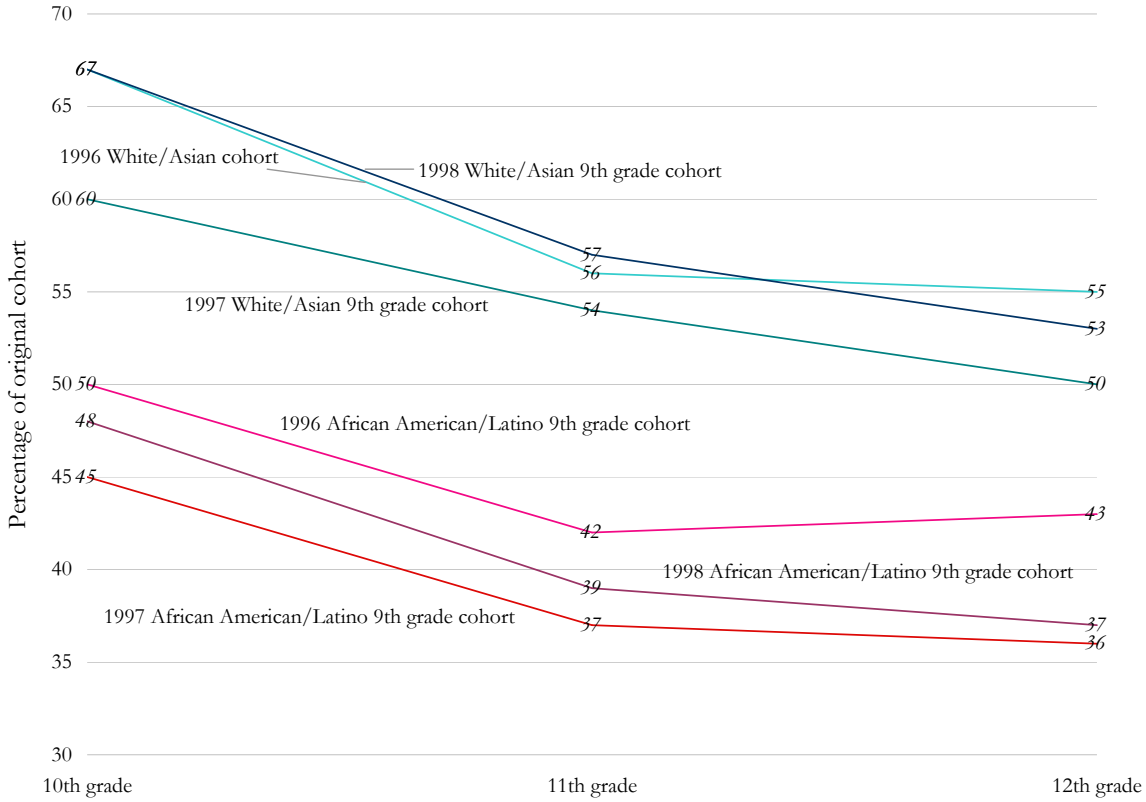


Figure 6. BCSD high school cohort progression by grade and race/ethnicity (percentage of entering cohort). Lines are divided by entering ninth-grade cohorts from 1996 to 1998.

Using a cohort method to track graduation rates for students entering 9th grade in 1997, Vasquez-Heilig (2006) calculated whether students were eligible for graduation based on having reached senior year and having passed of all sections of the TAAS exit exam. As Table 2 shows, of the 13,651 students in the 1997 9th grade cohort, 4,111 (30%) appear to be eligible for graduation within five years. This is close to the district-provided graduation status code, which shows that 4,458 students (33%) are coded as having graduated by 2002 (a five-year span).

Table 2
Graduation of 9th grade cohort entering high school in 1997

Student Characteristics	Graduation Eligible within 5 Years	District Reports Graduated within 5 Years
Overall	30.1%	32.7%
White	44.9%	43.3%
Latino	26.1%	24.8%
African American	29.3%	39.4%
Asian American	53.1%	49.4%
Economically disadvantaged	26.3%	28.3%
Limited English Proficient	14.1%	20.0%
Did not pass 8th Grade TAAS in reading	7.3%	19.3%
Did not pass 8th Grade TAAS in math	9.7%	22.3%

Asian Americans had the highest completion rate at almost 50% coded as graduated. Whites and African Americans had the next highest rates at 43% and 39%, respectively. Fewer than a quarter of Latino students in the cohort were coded as having graduated from BCSD by their fifth year. Students coded economically disadvantaged showed a graduation rate of approximately 28%, while only 20% of Limited English Proficient (LEP) students were coded as graduated. What happened to the 60% of minority students that did not graduate in five years? What happened to the 80% of LEP students who seemed to merely "drop in" to high school? Did the waiver policy that held students in 9th grade who did not pass their four core courses effectively push minority students out of school? The positive answer to this question is verified in the following case study.

A Case Study in Compliance

In this section of the paper we can confirm and extend the findings of the longitudinal dataset of 271,000 students in BSCD from 1995 to 2002 with a case study that shows a representative example of a school which experienced the unintended consequences of the accountability system. The researchers were evaluating the work of Edgeview High School for a major education reform project when they realized that over half of the student body was remaining in the 9th grade even as the school's high-stakes testing scores caused it to become a Recognized School making Exemplary Progress in BSCD. Edgeview was also one of the schools that reported 0% dropouts during 2002. The competing pressures of the accountability system's demand for high test scores and of the school personnel's attempts to provide a quality education for their students become painfully evident when we closely examine a school. Analysis of this school, illustrative of BSCD schools in the 1997–2001 period of the statistical section, revealed the full impact of the accountability system on the students and their educational experience, not just on their passing or failing the state test.

This case study of Edgeview High School is also valuable because it shows the avoidable losses of low-income and minority youth in a representative BCSD high school. We can also see the way the system affected specific groups of students, their teachers, and administrators. Edgeview High School was a reforming school, working to personalize the environment of its students with small schools when the effects of that district's demands to improve test scores intersected with the small schools reform. As the initial period of high-stakes testing took hold in BCSD, this high school was not bringing up its high-stakes test scores high enough or fast enough to meet district goals. For the school's leadership, this created a confusing dilemma: satisfy the accountability system

or satisfy their own sense of responsibility to provide a strong curriculum and help students remain in school until graduation. This set of demands wore on the two consecutive principals in such a way that each spoke and acted as of two minds when leading the school. The case study below will illustrate how this was manifest in developments at the school.

The accountability demands

In 1995, as Edgeview was restructuring the school into five small learning communities meant to increase personalization and student achievement (what it called "small schools" within the school), the school district was simultaneously implementing reforms to encourage its schools to make strong progress on standardized test scores. One of the school board members during this period repeated a commonly held theory regarding what the accountability system would do for schools: It would raise teacher capacity and result in the removal of poor teachers.

> The test was able to show the teachers, "Look, your fourth-period algebra class is not getting it. Do something." When that principal was able to move some of these people, or encourage them to retire, or maybe they saw that, you know, the fire was getting hot and decided to leave, and then the math scores improved.

According to this theory, high-stakes testing should have facilitated strong teaching and programs to hold high school students in class. Instead, the obsession with gaining high test scores forced teachers to abandon their regular curriculum to teach to the test and to seek other ways of boosting scores (McNeil, 2000, 2005).

Initially, Edgeview was not one of the high performing high schools on the standardized tests in BCSD. Its marginal position in the city and in its immediate community mirrored its position in the school district. However, the school demographics reflected those of most of the schools in the district. The ethnic profile of the student population showed a growth over time in Latinos and a decline in African Americans: over 75% Latino, less than 20 % African American, and less than 4% White. Slightly over 10% of its students were labeled Limited English Proficient (LEP) even though Spanish was the language commonly spoken in the halls among the students. Acknowledged as one of the poorest schools in the district and the state, this school received marginal funding, with virtually no allocations for science labs, its library, or student activities.

This low-performing, poorly-resourced, and high minority school did make modest gains in its test scores, 15–20% increases in math, reading, and writing in 1996–97, but the marks were not up to district standards. All of the first principal's evaluations were tied to the school's standardized test scores. The single line typed into the "Consequences if Goals Not Met" line on her performance evaluation for that year was one word: *termination*. Her job was on the line, and her school's progress on the 10th grade exit tests was not acceptable to this district.

The ninth-grade waiver

During a chance conversation with another principal in 1997, this troubled principal complained, "I don't know what I'm going to do about my TAAS scores. My school always ranks near the bottom of all the high schools in this district. I'm in danger of losing my job." With a hint of her former astonishment, she related in our interview his surprising reply:

> "Well, why don't you apply for a waiver?" I asked, "A waiver for what?" And he said, "Well, how do you think these other schools are getting [our scores] up there? We all got waivers."

The Edgeview principal then explained to the interviewer that test scores were reported back to the schools in the spring of each year. In the following fall, all of the principals received a waiver book from the state education agency listing the waivers for which each school had applied. The principal told us that she had seen a list of several schools (including the other schools in the ethnographic analysis in this study) that had applied for the waiver that her fellow principal had been telling her about. It was a waiver from the traditional method of basing grade promotion on the number of accrued credits; a school under this waiver could base grade promotion on different criteria, such as having to pass four core courses rather than gaining credits. As described by a teacher from Crockett High School (a school in this study),

> There's another policy: if you [a student] don't pass all of your core courses, you are not promoted to the next grade. So you may have 15 credits, but you may not have passed Algebra I, and you're still classified as a ninth grader.

The teachers and the principals in the district were aware that this was a policy causing students to be intentionally held back in 9th grade. The rhetoric surrounding this policy centered on "making sure students are ready for the state 10th grade test." But the sense among teachers and principals interviewed for this study was that the stricter criteria, which excluded credits in languages, the arts, physical education, and other electives from being counted toward grade promotion, held back weaker freshmen students to raise the school's aggregate scores on high-stakes tests. A review of several transcripts and accounts from a number of students revealed that some students had to wait as long as two years to be assigned to the subjects they had failed, thus further delaying their entry into the tested cohort. The administrator of one of the learning communities (or "small schools") at Edgeview explained,

> All the schools had it. In fact, we were one of the last schools to accept the waiver because philosophically we thought [to do so] would be cheating. But Fine Oak High School is… the other high school in our [area]. We were always having our noses rubbed in the fact that our scores were not as good as Fine Oak. Fine Oak wasn't testing everybody, but we were testing everybody, we felt we had no choice but to move to that waiver in order to save face and get our scores up.

The Edgeview principal and her staff felt "they were cheating" but that they had to yield to the pressure of the test and the demand for high scores.

The waiver had put the principal and her staff in an untenable position, with irreconcilable pressures. In one breath, the principal was saying that the teachers were proud of the small learning communities they had developed. As she marveled, "They really felt we were pioneers." This school was the only one in the district with a structure of five learning communities, or small schools inside the school. In the very next breath, the principal was explaining, "Ethically, I think it [the waiver] is wrong. But if I'm going to lose my job if my scores don't go up, do I roll over and forget it?" The principal realized that the students would and did suffer with the enactment of the waiver. As she explained later, the waiver identified the retained students as "losers" and pushed them to leave school. She explained, "A lot of the kids were defeated, particularly those in the weaker small schools. And they dropped out." The binary set of pressures faced by this principal is representative of the dilemma faced by other administrators in the district. She grappled with the predicament of holding her job at the expense and actual loss of the students this system was meant to serve.

The push of the accountability system for strong numbers split the identity of principal: Was her job raising test scores to produce high ratings or improving the school to meet the needs of its minority students? As the BSCD school board member claimed earlier, the expressed purpose of the push for high test scores and the strong growth in testing marks was the improvement of teaching and its reflection in student work. But as described in the previous section of the paper as well as in

the next, the effects created a gap between what was good for the state's reputation as a leader in school reform and what was good for the students and their teachers.

The power of the waiver that served the accountability system with its sanctions and rewards won out, as this embattled principal succumbed to pressure from the area office to use this safety valve. In doing so, she managed to raise her school's standardized test scores at the expense of the students, the avoidable loss of students. The principal was forced out of her position the next year, not because of the dropouts, but because even with their leaving, the school's scores did not rise fast enough to satisfy district and regional office objectives. Her successor continued the waiver, continued to retain large numbers of 9th graders. As a result, the second principal saw the strong progression of school test scores under the waiver. With the help of the waiver, Edgeview's accountability rating moved from "acceptable with acceptable progress" to "recognized with exemplary progress." This jump earned the school its exemplary progress rating on the district's accountability matrix and another star on the marquee in the front of the building for becoming a Recognized school. Scores on standardized tests rose substantially for the school years 1998–2001. But what price did the students of Edgeview pay for these strong high-stakes test scores and their accompanying reward of exemplary progress for the school? The statement of the principal who said that these minority students "were defeated" and "dropped out" was validated by the analysis of Vasquez-Heilig (2006), who showed the high school withdrawal rate for seven BCSD high schools from 1998–2001.

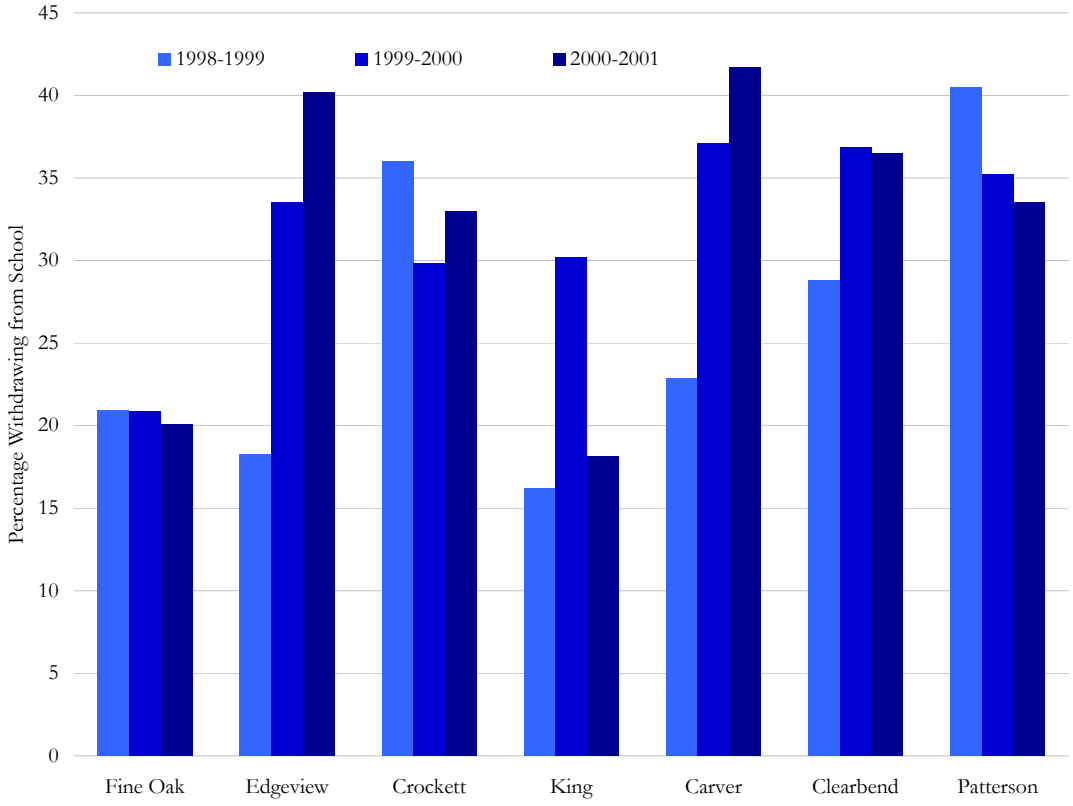


Figure 7. Withdrawal from high school: Seven schools in BCSD.

Notice that Edgeview entered the race for high scores late, so it did not join Crockett, Carver, Clearbend, and Patterson, schools using the waiver, in 1998–99. By the following year,

Edgeview had one of the highest withdrawal rates, rising from 18% of students in 1998–99 to 40% in 2000–01. The underside of the accountability system was exposed as the loss of minority students was intensified. Ironically, when 40% of students were withdrawing from school, Edgeview High School reported no dropouts. This discrepancy was revealed to the nation with similar reports in other high schools in the state (Schemo, 2003). The picture of high school withdrawal was matched by the number of students retained in the 9th grade. Vasquez-Heilig documents that increase in the table below.

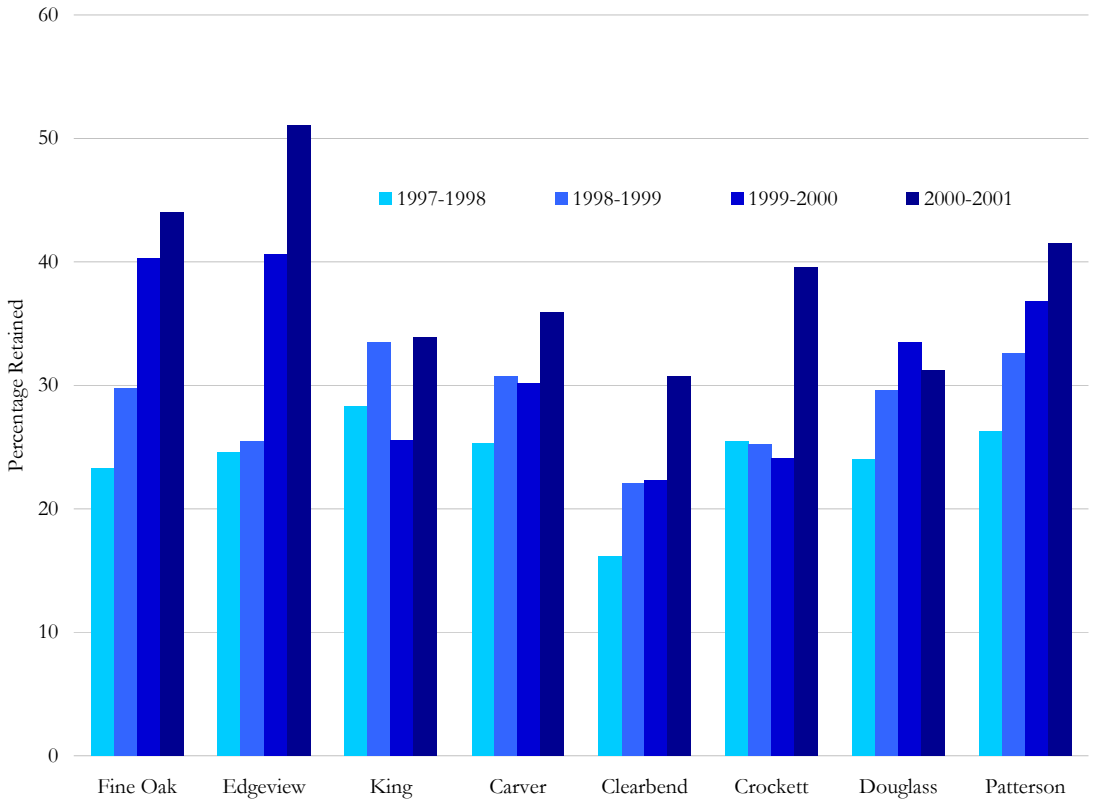


Figure 8. 9th grade retention: Eight high schools in BCSD.

This time, Edgeview has the distinction of being the school with the highest percentage of 9th grade retention after a moderate start in 1997–98, the first year that the school used the retention waiver. The retention rates grew from 25% in 1998–99 to over 50% in 2000–01. The new principal was appalled at the high retention rates for 9th graders. In a faculty meeting, she told the teachers that although she was opposed to the waiver, she could not end it for the high Texas Education Agency rating it brought the school. Again, the question of what was best for the students lost out to what was best for the school's standing in the district.

Dissolution of learning communities. This new principal of Edgeview believed that a 9th grade small school with no 9th grade repeaters would stave off the effects of the ninth grade glut created by the waiver. However, when the numbers came in, the freshman learning community had only slightly more than 50% passing on to 10th grade. Starting a special learning community for 9th graders and separating the repeating 9th graders could not forestall the 50% retention rate that was established with the waiver policy.

In the fall of 2001, the faculty expanded the 9th grade small learning community to include the passing 9th graders plus new 9th graders into a 9th–10th learning community. Again, the passing

rate only hit 50%. This second principal and the faculty had created the 9th grade small learning community to facilitate the 9th graders' transition into high school and on to the 10th grade, but the district hierarchy's directive to raise test scores took precedence over the school's reform efforts. The creation of the 9th grade small learning community and then the 9th/10th grade learning community forced the dissolution of the career-themed small learning communities. The two remaining small schools became Track Three and Track Four. Track Four served as the dumping ground for the retained 9th graders.

Dropbacks. The retained 9th graders were now called *dropbacks*, and their repeating classes were *dropback classes*. The students in these classes were representative of the ethnic and gender representation in the school. Some of their teachers were derisively called *dropback teachers*, and their treatment by consultants and teachers in other learning communities was negative. This naming of students, teachers, and classes replicated what Anagnostopoulos (2006) found in Chicago when 9th graders were retained for not passing the state's high-stakes test. Just as the students in the Anagnostopolous study had their identities set by their position in school, so these students and teachers knew and felt their positions as dropbacks. With no strong academic peers, these repeating 9th graders responded to their teachers by coming to class without books, writing instruments, or assignments. Verbal responses were jokes or sarcastic asides. This climate of failure was exacerbated by an us-versus-them standoff between students and teachers. The teachers saw the students as unmotivated, and the students saw the teachers as uncaring.

The repeating 9th grade student that Radigan (2003) shadowed between 1999 and 2001 in this disruptive dropback track showed the negative effects that labeling and positioning in school can have on academic work. He was unable to pass Algebra I until the summer after his *third freshman year*. After his third year of failure, the 9th grader said, "I'm failing algebra. I failed algebra three times. I'm terrible at it. I don't understand one thing."

For some students the label of dropback extended into the course scheduling system as repeating 9th graders were frozen into the 9th grade courses even if they passed one or two of their core courses. Two former students recalled specific examples of scheduling that kept them in freshman classes because they were labeled as ninth graders who did not pass all four of their core classes. The first student's position as a dropback made him a perpetual 9th grader in the system for any course he may have passed.

> Well, that last two years that I was in ninth grade, there were finally classes that I passed and got credit for. But they would put me in the same classes again, so then they would catch that later in the year, which they couldn't do nothing about it, you know.

Another dropback spoke about that the way his repeated placement in the Algebra class he had passed drove him out of school.

> Oh, yeah, they had me taking Algebra forever. I passed the first year, so in the second year I just decided not to go. I tried to get it [the course schedule] fixed, but they wouldn't fix it. So after the third week trying to get it fixed, I just stopped going.

Edgeview was paying a high price for holding students in 9th grade so that test scores would continue to rise. This school and its counterparts with poor and minority students in BCSD (as described in the statistical section above) were effectively retaining students in 9th grade for as long as three years, thus facilitating a 40% or larger push-out of avoidable student losses.

The unbridled power of the accountability system. The second principal of Edgeview found herself with contradictory identities. In one identity, she acquiesced to the demands of high-stakes testing with her role as accountability manager, retaining students in 9th grade for not passing their four core classes. In her oppositional role as champion of the children, this principal tried to

mitigate the greatest 9th grade retention rate in the district by creating a 9th/10th grade learning community, designed to keep students attending school. Edgeview gained a high accountability rating, but the dropback students who had been retained in 9th grade for as many as three years eventually became dropouts. In an interview with us, this principal, seemingly of two minds, decried the Texas Miracle and exposed the façade that hid the elimination of thousands of poor and minority students in BCSD.

> It's not a miracle to manipulate things. A miracle is saving kids actually, in reality—that's what miracles are. To go out and get these kids who were dropped out, or to get kids who are not achieving and find ways. That's a miracle to get all of it to do that. It's not to manipulate things so that it appears—it's a façade.

The façade of the accountability system made Edgeview seem like a school that was narrowing the achievement gap between majority and minority students. With the high test scores, the reported 0% dropout rate, the exemplary progress, and recognized school status, this was a strong part of the Texas Miracle deconstructed in the previous section. However, the realities of the accountability system encouraged a 9th grade retention rate of more than 50% and over 40% student withdrawal rate.

The dilemma *to comply or to educate* produced the dual identities we witnessed in these two principals. The pressure to comply with the accountability system came with the desire for this Texas Miracle. This school, representative of the district, with 75% Latino and 20% African American population, needed to advance its accountability rating from Recognized to Exemplary. Holding 50% of these minority students in 9th grade as liabilities accomplished this mission. The hierarchical nature of this process was a taken-for-granted norm to the two principals in this study. The conditions under which they worked have not changed with the movement of the tests from TAAS to TAKS. The area office is expected to move all the schools in that subdistrict to a higher level. And in turn the superintendent must push more schools in the district up to a higher level if he or she is to earn a huge cash bonus, else possibly lose his or her job or be put on probation. So the signal to produce higher accountability indicators comes from the state to the superintendent, from the superintendent to the area superintendent, and down to the principal. This message is communicated through city-wide principal meetings, central office memos, banners and press, signs outside the schools touting the school's ratings, commitment of dollars to test preparation materials even when dollars are scarce for curricular materials and supplies, district commitment to test vendors and test consultants, and a large bureaucracy to oversee all of that—including adding test coaches in the schools.

Where was the hierarchical support for the mission to educate these minority youth in an under-resourced school with a test-driven faculty? The attempts to maintain a small-schools structure with adequate numbers of teachers by both principals garnered no district support. The attempts to lower the number of 9th grade retainees and stop the push-out of students by the second principal were neither recognized nor encouraged by the district. It was the mission of the accountability system that triumphed in the form of the Texas Miracle.

Tracing the pressure to comply with a highly standardized accountability system over several years in an urban high school shows that even a well-meaning faculty eager to improve the school can be rendered ineffectual by the short-term pressures to produce numbers. The Edgeview case study shows how attempts at real school improvement were less likely to be rewarded and sustained than those practices which resulted in higher dropout rates among its students. In BCSD, reforms which require time and resources were discarded in favor of raising school scores. When the most powerful, most dominant, and pervasive policy is an accountability system that ranks schools on the basis of student test scores, real measures taken toward structural reform are thwarted. And as Edgeview High School demonstrates, without a serious investment in the educational capacity of the

school, pressure to produce indicators can lead to school practices which subordinate educational quality and students' progress toward graduation in favor of the production of test scores. When the leadership and faculty jobs are riding on student test scores, the students themselves can come to be seen as assets or liabilities to the schools' ratings. The Edgeview case study shows how dropping out is, in many cases, a result of the accountability system when it is working as it was designed. The waiver that produced such high 9th grade retention rates and stopped 60% of the students from graduating is now normalized into the accountability system as a customary and expected practice. The findings at this school were confirmed in the other schools we studied, that each of the other schools faced problems in maintaining high test scores while avoiding losses of its impoverished and minority students.

Students in a Zero-Tolerance Environment

Our case study of Edgeview showed us how the waiver policy in the accountability system worked in practice for a school that was trying to comply with test-based accountability ratings while undertaking reform. To investigate whether the pattern of compliance and dropouts at Edgeview was an exception or representative, we expanded our qualitative study to look into other high schools and middle schools in the district. We were allowed into the schools to ask how students were experiencing the system. Our statistical analysis and case study encompasses a period in which the accountability system had been fully operationalized for several years. To ascertain whether our findings persist to the present, we returned in 2006 to the schools in the sample to interview students, teachers, and administrators.[5] From these interviews, we learned that students continue to live under essentially the same system, which is now normalized.

Here we analyze how the workings of the accountability system and dropping out of school play out for and are understood by students in a representative sample of BCSD schools. We used focus groups, interviews, school-level observations, and analysis of documents to explore these questions. This qualitative component is based on a data set of 122 students and 38 administrators and teachers in seven different schools, including our case study school. Through analysis of these data, we concluded that while students face a great many obstacles to school completion that have already been well documented (e.g., Ekstrom, Goertz, Pollack, & Rock, 1986; Orfield, 2004; Rumberger, 1995, 2000, 2004; Wehlage & Rutter, 1986), the accountability system adds additional, avoidable factors to the mix, thereby exacerbating the problem. In *Framing Dropouts* (1991), Fine documented that schools may behave in ways that promote rather than prevent dropping out. Educators, policymakers, and concerned citizens often ask what is "the cause" of such a high dropout rate. If we, as researchers, could point to a single cause, this problem would be far easier to fix. This study demonstrates that the accountability system aggravates the high dropout rates both through its incentive structure for adults and through its effects on students. We now understand the causes of dropping out in this policy context as both *additive* and *convergent*.

We have come to understand the causes of low graduation rate as *additive*. Although any individual cause may only lead to a small proportion of dropouts, when added together these causes lead to the low graduation rates we have found. Considerable prior research has found that dropping out is associated with such factors as income level, disengagement from or frustration with school, family demands, pregnancy, interest in gangs, or illegal methods of making money. The results of our qualitative research demonstrated that the accountability system adds further to these causes by

[5] We acknowledge the participation of Elle Rustique-Forrester on the research team that conducted these interviews.

creating an environment that is highly punitive and where adults at the school level experience strong incentives to simply allow students to exit the system.

From interviewing current students, both those who have remained in school as well as former dropouts who have returned, we can begin to understand that the factors causing students to drop out are also *convergent* upon each individual student. A student's decision not to return to school is not defined by what researchers or policymakers would consider a single cause. In reality, students experience several of these simultaneously, until they converge so strongly that staying out of school seems a better decision than staying in. In traditional policy analysis, effects are tracked from the top down. In this analysis, the voices of the students led us to conceptualize an accountability policy based on the lived experiences of the students. Through this study, we learned that the accountability system adds further to the negative factors for students by creating an alienating environment that is both bureaucratic and punitive.

Through our interviews and focus groups at seven representative high schools in Brazos City, we found several components of the accountability system that youth tell us make it increasingly difficult for them to persist in school, adding to the other factors that already existed. These factors all arise from systemic accountability policies that surround the students and powerfully shape the school environment with which they interact on a daily basis. Although high-stakes testing is commonly seen as a separate component of accountability policies, analysis of this data forced us to rethink this compartmentalization. We came to understand that the testing is one part of a cluster of policies and practices with a common thread: to create an appearance of bureaucratic control and toughness while raising school ratings.

Just as the waiver policy interacted with the work of teachers and students at Edgeview High School, the case-study school, we found that this cluster of policies reinforced one another in ways that created additional factors related to dropping out, and converged upon the students, on top of the pressures already incumbent upon them. And, as demonstrated in the quantitative analysis above, this problem disproportionately impacts minority and poor students who are already at heightened risk for dropping out.

Below, we describe in more detail the three factors related to the accountability system that we identified through this study. First, high-stakes testing and accountability strongly influenced the curriculum in ways that were increasingly test driven, leading to fragmentation and causing even more disengagement by students than existed previously. Second, the widespread practice of retention that held students in 9th grade without having to take the 10th grade TAAS exam led to a great deal of confusion and frustration as students found they could not advance through their high school grades in a normal fashion. And third, zero-tolerance attendance policies that policymakers conceive as separate from the accountability system in fact were connected to the overall bureaucratic and punitive culture of that system. They create disincentives for students to remain in school by withholding course credit after only three unexcused absences and push students with frequent absences into the court system, essentially criminalizing the absent students with little regard to the reasons they miss school.

Degradation of curriculum

A theory behind basing teacher and administrator rewards on measurable indicators of student learning was that it would pressure teachers to improve instruction and, presumably, the quality of what is taught. High-stakes testing and accountability was shown in these schools to have, over time, significant effects on the enacted curriculum (DeBray, Parson, & Avila, 2003; Firestone, et al., 2002; McNeil, 2001; Sloan, 2005), effects discernible to students as less than positive. In some

cases, the curriculum is degraded to simple test prep: drills in the form of the standardized tests in the guise of practice. However, even where this does not happen, the curriculum drifts towards ensuring that students can answer the questions that will appear on the tests, so that the form of the tests becomes the de facto curriculum. This narrower, more rigid curriculum affects students and their motivation to complete school. How does the specter of high-stakes testing and its consequences—which hangs over students, teachers, and school staff—affect actual instruction, and how, in turn, do students respond to their experience with that instruction?

Students who dropped out and returned to school and those who were struggling to stay in school all reported that high-stakes testing drove much of the curriculum. In many cases, this translated to a good deal of direct test preparation, either specific courses to prepare for tests or class time devoted to direct test-preparation activities. One focus group of students at Lincoln High School were asked if they read books in English classes or studied only what would be on the TAKS. They agreed that the ratio between direct test preparation and other activities in their classrooms was "about fifty-fifty." Reports from teachers confirmed that high-stakes testing controlled the curriculum.

Students and teachers in this study told us that significant time and energy were spent in their classrooms preparing directly for the test. These students are sharply aware that the content of the high-stakes test is the focus of their curriculum and should be the focus of their learning. As one Crockett student said, "If you're a freshman, you practice it all the way through until you take it in your junior year" [the exit-level test, formerly administered to 10th graders, is now in the junior year]. The Crockett students reported practicing for the test in all of their classes, even gym. Linea added, "They had specific classes for the TAKS, like, TAKS Writing. and TAKS Math that we had to take." The TAKS is a computer-scored, multiple-choice test, and practice often meant spending class time working through multiple-choice practice questions rather than comprehensively studying the subject at hand.

A focus group of African American honor students at Lincoln felt that their classroom experiences had changed markedly since their teachers began paying increased attention to test scores. At the time of the interview, this particular school had been targeted for possible closure the next year due to failure under the Texas and federal (NCLB) systems of accountability. This group of honors level, college-bound students were able to describe the aggravation they felt when the focus of their academic curriculum was changed to test preparation. In one student's words,

> Instead of teaching us the real life things that we are going to need for college and stuff, they started zeroing in just on that test. So it makes everybody nervous, and it threw everybody off. So, like, our curriculum is thrown off, 'cause what they originally were teaching us in the subjects, all of the sudden they switched, and then they were just zeroing into this test.

Paradoxically, the students said that despite so much focus on the test, many students did not pass, leading to even further frustration. Lenicia claimed to be "over-educated for the test." When the test day came, and many students didn't pass, the sense of personal failure was huge. For these students the whole test-preparation process hindered rather than helped them in their pursuit of an education that would prepare them for college.

Students also perceived a shift in relationships with their teachers as the focus in the classroom moved away from the students themselves and toward the test. An African-American senior in the honors program explained,

> [I've] been here four years. The first two years were great to me. I learned a lot because it was so—you know—we had a relationship with our teachers….And, like, as soon as they brought TAKS in for real, like you couldn't graduate without it, all the teachers started getting so scared because they want us to succeed at it.

She elaborated, "Some teachers are so scared and don't know what to expect on the test, that they zero in on that test and it bugs us just hearing about this test." Tyrell, another honors student, revealed that he understood how high the stakes were for the teachers themselves:

> Where you learn so much and then even if the teacher does what they can to teach that student, when they [the students] get to the test they fail—that teacher might get fired. They make sure that they drill you, drill you, drill you, and then….

There is evidence that the narrowing of the curriculum in response to test mandates further widens the inequities between poor and minority students and their more privileged peers. Under NCLB, more grade levels come under the high-stakes testing requirement. Gold (2007) found that this formula narrows the curriculum for minority students in urban schools, while the suburban schools retain a broader curriculum because they are not under a similar pressure to make a large increase in test scores. Fine (2005) finds that suburban students in Advanced Placement courses are offered a broad curriculum, whereas the minority students in New York City are offered a narrow, test-prep curriculum under the increased emphasis on high-stakes testing at each grade level.

When we interviewed students, they often reported that the work was not "broken down," meaning that they could not understand it well. They expected the teacher to help them understand the material and were disappointed and frustrated when that didn't occur. As one Edgeview student explained, "Some of the teachers out there, they just put the work on the board and tell us to do, like, do it." With no explanations, the students are often expected "to know the material" or "to learn it" on their own. The high-stakes testing and accountability system does not necessarily translate into increased instructional capacity. The students and teachers are pressured to do more, and do it better, yet they struggle with the teaching and learning of the material. One student explained, "We have to memorize everything they're teaching us, which is really hard." A lack of teacher capacity to work toward conceptual learning leads students to believe that they must memorize the practice tests they are handed to prepare for the high-stakes test. Some teachers hide their own knowledge of their subjects to just do "drill and kill." For students already on the academic margins, this leads to frustrations so great that they begin to turn away from school altogether.

For these students, the high-stakes testing and accountability system did not lead to any improvement in the instruction offered them. Rather, it directed teachers' attention to raising test scores such that that the form of the tests and the material that would appear on it became a paramount focus in the classroom. Such focus makes the material more rote, less engaging, and more high-pressured for students who are struggling to stay in school and succeed there. The system was working as designed, with teachers responding to its mandate to produce passing test scores, but perhaps not as intended, in terms of improving learning.

Increased grade retention

From the case study of Edgeview High School and the quantitative statistics, we learned that grade retention was a strategy used by adults to assure positive ratings in the accountability system. Ninth-grade retention of low-performing students increased the test scores of minority schools and minority subgroups in multi-ethnic schools. These scores were essential to schools' earning a higher TEA accountability rating category. Schools were provided this "safety valve" for raising test scores, and most of them took advantage of it. A sanguine view of the relationship between retention and the accountability system might be that students need to repeat a grade to provide them more time

to get ready to take the exam. However, neither the adults nor students with whom we spoke ever described the process in this way, and we know from the findings of this study in combination with other research cited earlier that grade retention contributes to the low level of school completion we observed. The students directly affected by ninth grade retention were frighteningly unaware of the severity of what was happening to them until it was too late. The students we interviewed had not grasped exactly what they needed to do to progress smoothly through high school in four years, nor did they realize they could be forced to repeat an entire grade, including courses they had already passed the year before.

Despite the intuitive expectation that student performance will improve if struggling students are remediated by grade repetition, most studies find negative effects on student achievement from retention in grade. Retention can increase the odds of dropping out by as much as 250 percent above those of similar students who were not retained (Rumberger & Larson, 1998). Reynolds (1992) and McCoy and Reynolds (1999) investigated the effect of retention and school achievement on low-income urban minority children and found retention had substantially negative effects on cognitive achievement in reading and mathematics. Summarizing several decades of research, the National Research Council concluded that low-performing students who are held back do less well academically, and are far likelier to drop out, than comparable students who are promoted (Heubert & Hauser, 1999). The Chicago Consortium for School Research found that although some students' scores improved in response to Chicago's high-stakes testing policy tied to grade promotion, the scores of low-scoring students who were retained declined relative to similar achieving students who had been promoted, and their dropout rates increased (Roderick, Bryk, Jacob, Easton, & Allensworth, 1999). Retention policies appear to affect poor and minority students disproportionately. A Texas Education Agency report confirmed that African American and Latino students as well as economically disadvantaged students were retained in all grades at a substantially higher rate than Whites or non-economically disadvantaged students in Texas (TEA, 2001). Denton's (2001) national analysis of state-level promotion policies and retention rates confirms that retention policies disproportionately impact low-SES and minority students.

On the surface, educators and the public tend to accept repeating a grade as an acceptable academic treatment for struggling students. Yet for the student who repeats a grade, the experience is so profoundly disappointing and discouraging that it contributes to a sense in the student that continued schooling may be futile (Roderick & Camburn, 1999). Students who are required to repeat a grade lose face in front of their parents, community, and peers. As the student grows older, she or he begins to feel increasingly disconnected from school and begins finding a more positive source of identity elsewhere. Such students begin to consider the other options available to them, such as employment or commerce on the black market. Ninth-grade retention is likely to yield particularly negative outcomes: More students drop out in 9th and 10th grades than at any other time during high school (Balfanz & Legers, 2004). A question could be raised as to whether administering high-stakes tests in more grades, including 9th grade, could prevent the use of grade retention as a mechanism for skewing a school's accountability scores. From our more recent observations in schools, we have learned that increased frequency of testing, and testing in more subjects, dulls students to the testing process and diminishes the seriousness with which they regard tests prior to the 11th grade exit test. In Brazos City, as our statistics have shown, the pattern of dropping out was exacerbated, by being retained in 9th grade, a practice directly resulting from pressures on principals to show increasing scores on their school's 10th grade test.

Students did not generally see the link between the accountability system and grade retention. Instead, administrators and teachers explained to us the connection between the accountability system's administrative incentives and retaining as 9th graders students not expected to pass the 10th grade test. For many students, this logic remained opaque; they knew only that

whether they were promoted seemed arbitrary and confusing. One perceptive teacher at Clearbend High School described the effect of the waiver on a voiceless student who did not understand how the system worked.

> And it really is a hard system to navigate through, and it really is hard to find out that information. And if you're already, you know, the kind of kid, which most of ours are, where you don't trust adults, and you don't trust somebody who is from a different ethnic background than you, then you're not even going to come ask why. You're just going to leave.

This teacher also understood the motivational effects on students who are held back and can't imagine getting only "two credits" for a year of work:

> They're not coming back. I mean, and it could be a kid that would come back if they [could say] "I'm a tenth grader" instead of "I'm a ninth grader."

These students enter high school at 14, generally, and are not clear on the consequences of their actions, particularly missing school or class. Many parents also have little understanding of how the system works, so that parents and students both may be blindsided at the end of the year when the news comes that a student will repeat ninth grade. They understand little about course credits and even less about the waiver system that allows a school to block promotion without course credit in each of the four core subjects. We interviewed some Latina students who as 9th graders had not been assigned to the four core subjects and thus could not have been promoted to 10th grade even with good grades in all their classes. A teacher at King High School explained that some of the confusion for parents and students comes with the transition from middle school to high school and the belief that students "pass a grade," not courses. This teacher has had parents say, "they have been here for four years, why not graduate?"

For the former dropouts with whom we spoke, grade retention seemed a mystery, something that resulted mostly from poor attendance, which led to course failure, which led in turn to being held back in the ninth grade. Students in a focus group of returned dropouts from Edgeview did not realize they were retained in 9th grade until they received their report cards. The full effect hit at the beginning of the next year when "we still had ninth grade classes, same teachers." Only then did these students realize that not getting enough credits meant retention in a grade, and they felt "stupid." No one in the school, neither their teachers nor a counselor, had explained to them and their parents why they were being retained or what would be needed to assure their path to graduation.

This exchange among boys at Edgeview High who had dropped out and later returned to complete high school followed an interviewer's query:

> Interviewer: So have all you guys spent time in 9th grade?….Did you have just one year in 9th grade?
>
> *Alberto*: About four.
>
> *Ernesto*: :Yeah, I did about four.
>
> *Cleavon*: I did—let me see. I had two years in 9th grade….
>
> *Interviewer*: Ernesto, you had four years in 9th grade?
>
> *Ernesto*: Three or four, I don't know….I dropped out twice.

Former dropouts from Edgeview spent "three or four years in 9th grade" with teachers who saw them as kids "who were not going to do the work." So these students dropped out. Some were not offered the courses they needed to move beyond 9th grade. Notably, at all the schools we studied, there were students who persisted. Yet, at several schools they were far outnumbered by those who left before ever reaching 10th grade.

The students most at risk of grade retention and dropping out are generally struggling with their academic work to begin with, so it is unsurprising that they continued to struggle, even failing

some courses. These multiple levels of discouragement and frustration are added to other forces in the students' lives that make it difficult to complete school. In end, if it is too much to handle, they merely disappear from formal education, dropping out of school. For large numbers of students, as we have seen, these are the unfortunate effects of an accountability system that provides incentives for massive grade retention because of retention's positive effect in school-level scores.

Zero-tolerance attendance policy

One feature of attendance policy in Texas places additional punitive pressure on students that compounds the retention waiver. Mandates for student attendance are specified in state-level legislation. A clause known as "the 90% rule" states that in order to receive credit for a course of study, a student must be present 90% of the days that classes are held within that course, or else must have his or her absences excused (BCSD website). For high school students, this 10% leeway generally translates to four possible unexcused absences within a one-semester course. On the surface, strict policies on student attendance may seem to reflect the common sense that students be present for classes or provide an excuse from parents. However, in practice, these policies are so overly punitive and bureaucratically applied that they perversely create barriers for truant students who seek to return to school. Students with even a small number of unexcused full-day absences lose credit for all their courses.

Family life and school attendance: For students struggling on the edge of staying in school, what may seem sensible becomes burdensome. Students who work outside of school to support their families may easily become exhausted and miss school. We interviewed students whose absences were caused by needing to assist with child care or the care of grandparents, often unpredictably. In cases where parents may have little education, be unable to write in English, or be unavailable because of their work, notes to the school may not be forthcoming even when the absence might qualify as excused. High school students who live with grandparents or siblings, or are emancipated minors, are often unable to produce a note. In schools where only a note from a doctor will suffice to excuse an absence, poor families who do not have doctors, or who rarely see doctors for minor illnesses, may understand the policy but be unable to produce a note.

A Clearbend teacher expressed a complaint we heard across our sample of schools, that the attendance procedures were part of "a confusing system." She described one of her students who had As and Bs in all of her classes but who lost credit because of attendance problems. The student's mother did not speak English and could not explain her concern when she called the school. The teacher also attempted to find out why, but could not discover the reason. She threw up her hands in frustration, as she told us:

> I think that—I mean, I feel that if I work here at the school, you know, I—English is my native language and so there's not a communication problem. I have no problem asking questions of people, and I still don't know what the attendance policy is at this school. If I'm frustrated with it, I'm sure that a student who is fourteen years old is frustrated with it. And probably their parents who aren't even on the campus and can't—that they're frustrated with it as well.

This teacher's frustration highlights not only the arbitrariness of the system but the fact teachers are often powerless to be effective advocates on behalf of students when policies are made and enforced at a level so removed from the teaching faculty.

Truancy and ticketing. The zero tolerance attendance policies do not end with in-school sanctions. Truancy policies also involve the criminalization of school absences. Texas law creates grounds for prosecution for truancy if a student is absent without parent permission in a way that

the state defines as persistent—three days within four weeks or 10 days within six months. Students with a string of unexcused absences may be catapulted into a court system, which terrifies them and their guardians. Worried about the prospect of appearing in court, or unable to pay fines, they will often withdraw from school as a defensive measure. This regulation shifts youth from the care and responsibility of the school into the juvenile court system. While this policy may have seemed sensible in conference and legislative hearings as a way to promote attendance, and may motivate some few students to attend more regularly, it creates a paradox as well. For a student who makes the mistake of missing a few too many days of school, or who does not understand the vital distinction between *excused* and *unexcused* absences, or even those who do not know how to explain family emergencies to school officials, the policy serves to alienate students to the extent that they may quit school. In our interviews, we learned that many of the English Language Learners (ELLs) miss school to babysit siblings and to translate at doctors' offices or care for elderly relatives, (Radigan, 2005), acting responsibly in their families while being seen as non-serious about school. Some of these students may intend to start clean after the end of the semester, but after beginning to create a life without school, they find it difficult to return.

The experience that many students have of law enforcement in response to absences is the receipt of tickets, with fines that can escalate to several hundred dollars. The theory behind law-enforcement action is that getting tough on the students and their parents will motivate more students to come to school. Minority honor students and low-performing students both suffered from these policies. In a focus group interview at Lincoln High School, some students from the honors program described their experience with this policy, demonstrating how being subjected to it feels to a student—even a good student. An 18-year-old senior student, Lenicia, explained that she had "so may unexcused absences" that the school wanted "to kick her out." Lenicia received "a truancy letter" saying that she "would be fined $500 a day for any further absences, and a court date would be set." Her friend Keisha, reiterated that the $500 fine is "per absence." Lenicia's reaction to this dilemma was "I might as well drop out." Tyrell, who was sitting next to Lenicia, echoed, "*I* might as well drop out." To solidify the ridiculousness of the situation, Lenicia added, "And no one's going to pay no $500 a day," as if she or anyone she knew would even have that kind of money. This student managed to get her absences approved, but added,

> Most of them [other students] get [kicked out]. And they're not telling you that you're kicked out of school either, so if you were absent that day, they'll send somebody around to all your teachers and the teacher would just sign your paper—your check-out papers. And when you come back to school they'll tell you you're not—you're no longer enrolled in school.

A high school student from Edgeview who had dropped out and recently returned to be part of a recovery program was one of those low-performing students whose retention in 9th grade made him a product of the waiver policy and facilitated his poor attendance and the accompanying tickets.

> Well, I flunked ninth grade and back to ninth grade again. Kept going in circles and circles, and my mom used to get tickets and tickets. So, you know, it's just left for me to drop out instead of me just giving my mom nothing but tickets.

This ticketing due to attendance problems served to exacerbate the students' sense that they did not matter, that they were merely numbers—assets or liabilities—in a high-stakes testing and accountability system. This was a ticket to court, a ticket out of the jurisdiction of the teachers and the school. The students were being ticketed out of school, and the absence of these low-performing students was raising the school's high-stakes test scores even as the dropout rate increased. At the same time, our analysis of the accountability system tells us that the adults in

the system have little incentive to track down these students, as they are most often the lowest-performing, the ones likely to bring down the school's average test scores if they return.

Converging effects. The accountability system is not fully understood by the students, but in interview after interview, at school after school, they were able to articulate the ways their lives and those of their peers no longer in schools are made more difficult and less educationally engaging by policies that make it difficult for them to persist in school. Each of these alone has serious consequences for youth, but as experienced by young people in these zero tolerance settings, each multiplies and magnifies the potential negative impact of the other. A test-driven curriculum, delivered by teachers pressured to focus on the test instead of their students and what they themselves know of their subject, becomes boring, not likely to entice students otherwise pressured to take on responsibilities outside of school, not likely to help them to manage to pass the tests with ease. Absences, whether due to disengagement from school or family and work obligations, set students up for punitive measures, including expensive court fines and even being withdrawn from school without their knowledge. Both of these—a degraded curriculum and zero-tolerance absence policies—stack the deck against succeeding academically, making students vulnerable to retention in grade, increasing the likelihood they will not receive the courses they need to move through school to graduation. Removed from the tested cohort, they inadvertently influence their school's rising accountability ratings.

Factors Influencing School Ratings in BCSD

To complete our qualitative and quantitative analysis and to corroborate or disconfirm the explanatory power of the accountability system's effect on dropping out, we looked at quantitative relationships between state ratings of school, on the one hand, and indicators of 9th grade retention and student attrition and other school-level variables, on the other Using multinomial logistic regression, we examined how school strategies translated into changes in TEA accountability rankings by conducting an analysis of the final years of the first-generation of Texas-style accountability (1997–2002). The independent variables used in this analysis included changes to the data described above (percentage increases or decreases in 9th grade retention, withdrawals according to the TEA student database, the official dropout rate, and students who disappeared from school but for whom there is no code in the TEA student database). Variables also included changes in other school-level indicators: demographic variables (percentage White, LEP, receiving special education services, or labeled "at risk") and changes in the indicators of school capacity (teacher turnover and the percentages of teachers who were novices and teachers who were fully certified). The dependent variable was an ordinal variable indicating the change in school ratings (same, rising, or falling), with identical ratings as the omitted category. For each school, a pair of consecutive years contributed one record to the analysis, with the change in the school-level data used as the independent variables. The independent variables were entered as groups in different models; Table 3 shows the results of four models, with the full or saturated model (D) presented last. The relationship between individual variables and changes in school ratings is presented as odds ratios, or relative changes in odds for the second TEA rating being different from the first.

The Bayesian information criterion (BIC) is utilized as the statistical criterion for model selection and fit for the multinomial regressions (Schwarz, 1978). The full model (Model D) has the lower value of BIC and is preferred over the models that enter student progress, school capacity, and student demographics as blocks. Model D shows that manipulating student populations, more specifically escalating 9th grade retention, strongly predicts better TEA ratings. The odds of a one-percent increase in 9th grade retention in a school that increased its TEA rating was about 24%

greater than in a high school with a TEA rating decrease, both before and after changes in student characteristics and school capacity are controlled. Additionally, as called for in the accountability system, officially reported 9th grade dropout rates show a negative coefficient in schools whose TEA rating rose as compared to those where ratings declined. As we have seen in the introduction to this article, these officially reported dropout codes bore little relationship to actual school leaving for students.

Table 3.
Multinomial Logistic Regression of TEA Accountability Changes (1997–2002): Odds Ratios and Bayesian Information Criterion

	Model A		Model B		Model C		Model D	
Variable	Falling	Rising	Falling	Rising	Falling	Rising	Falling	Rising
School Capacity								
Δ % Fully certified	0.13	1.05	—	—	—	—	1.09	1.28*
Δ % Novice teacher	0.77	0.96	—	—	—	—	0.85	0.92
Δ % Teacher turnover	1.05	1.02	—	—	—	—	1.11	1.08
9th Student Progress								
Δ % 9th disappearance	—	—	0.89	1.00	—	—	0.94	1.14
Δ % 9th retained	—	—	1.24**	1.34**	—	—	1.25	1.55*
Δ % 9th withdrawal	—	—	0.05	0.94	—	—	0.86	1.05
Δ % 9th dropout	—	—	0.86	0.52*	—	—	0.65	0.16**
School Demographics								
Δ % White	—	—	—	—	1.22	1.00	1.37	1.97
Δ % LEP	—	—	—	—	0.87	0.77	0.83	0.85
Δ % Special education	—	—	—	—	0.85	0.53	1.03	0.30*
Δ % At-risk	—	—	—	—	1.22**	1.04	1.21*	0.96
N	11	16	12	16	12	16	12	16
BIC	-23.9		-9.3		-5.9		-35.3	

*$p < .05$ **$p < .01$ (See Appendix B for standard errors and coefficients.) Delta signs (Δ) indicate that each independent variable represents percentage change.

The multinomial regressions also control for the impact of changes in teacher quality and student demographics on accountability ratings. We found that while exerting a somewhat small influence, in the full model the schools that experienced a rise in their accountability rating were significantly more likely to have an increase in the number of teachers who were fully certified, an important policy variable in a district with large proportions of uncertified teachers that are disproportionately allocated to poor and minority students. Other changes in teacher capacity (novice teachers and teacher turnover) did not show significance for higher TEA ratings. Finally, schools that had stable ratings were more likely than others to have experienced an increase in students classified as *at-risk* in models that already controlled for race, language, and special education status, while schools that increased their ratings were more likely to have lost special education students than those whose ratings decreased.

Together, these findings suggest that low-achieving students were significantly less likely to be included in the calculation of high-stakes accountability ratings, in addition to being more likely to have under-prepared teachers and to be in schools with a greater proportion of at-risk students. The results are compatible with expectations, given the accountability system incentives described by BCSD school staff throughout the paper. In addition, findings regarding the effects of 9th grade

retention on accountability ratings support the contention made by Holmes (2006) that when large numbers of students are retained in a grade, the next grade's scores are higher because the low scorers are removed from the testing pool. Although school and district scores and, therefore, accountability ratings may go up, the long-term consequences for individual students are invisible in the yearly snapshot of high-stakes accountability rating reporting. This invisibility is exacerbated by the state's leaver codes, which in effect, counter the official accountability policy's inclusion of a requirement to attain low dropout rates. Only the disappearance of a very large proportion of students made possible the Texas Miracle, a mirage of rising test scores and accountability ratings shown earlier in Figure 1.

Legitimating Avoidable Losses

The study reported here is the first to track the impact of high-stakes accountability on the youth in a way that combines qualitative and quantitative data. It documents that a direct effect of high-stakes accountability is the loss of thousands of youth, particularly students of color, from our public schools. These losses are intimately tied to behaviors that Texas's accountability system stimulated and maintained. Our analysis of a large cohort of youth in a large urban district over a multi-year period shows not only which young people are being lost from the system, but in which years. Being able to track individual children through the grades shifted the analysis of the policy in two ways. First, it permitted the research team to focus on individual children, thus stepping outside the accountability system's unit of analysis, the school. Second, it showed not just the binary of *dropping out* or *graduating*, but the point at which students left school.

The same data show that school ratings—the accountability system's key indicator for success—improve in proportion to these losses of low-performing students. That is, the system is able to produce positive indicators precisely because weaker students are being triaged out of the system before they, in the form of their test scores, can become liabilities to the school's publicly reported ratings. And this triaging is not an unknown or unwitting process. It is in complying with the system's pressure to produce rising test scores and rising school ratings that school staffs are making the decisions that result in increased dropouts at their schools. The state's system of leaver codes, another official policy component, obscures the scale of these losses.

Those decisions are consistent with the requirements central to the accountability system. Thus, the system's own incentives are complicit. The statistical analysis of the 271,000-student cohort over a seven-year period reported here documents the pattern of student losses. The case study uncovers the compliance mechanisms at work: the conscious decisions principals and school staffs make as they try to balance genuine school improvement and compliance with the accountability system. In interviews at other schools, we confirmed the fact that this school is representative not only of the costs of compliance but of the school staff's awareness of those likely costs. The pervasive patterns across a major urban district and the large student population sample raise the question of whether similar anomalies between rising test scores and rising dropouts in the state's other urban districts can be explained by the dynamics we have uncovered in Brazos City between the accountability system's incentives and the losses of youth from our schools, a number that in Texas reaches 135,000 each year (Johnson, 2007). The incentives in the accountability system tilt administrators' decisions toward compliance, although they know their focus on school ratings and test scores will succeed only when their weaker students exit the school.

This study would have great significance for our understanding of educational policy and models of centralized accountability if its only contribution were documenting the link between compliance with the production of rising test scores and the premature exit of minority youth from

our schools. But the study's focus on youth and their experience under such a standardized system reveals that a convergence of policies built into the accountability system exacerbates the pressure on youth and stacks the deck against persistence in school for many youth, particularly those who are poor, immigrant, English-language learning, African American, or Latino.

Students understand the accountability system as pressure to pass the state test to graduate and to protect their teachers' jobs or even, in one case, the future of their high school. The student interviews reveal that they experience this pressure as courses that became little more than test practice, confusing and opaque rules that resulted in their having to repeat 9th grade, and attendance policies that at worst sent them into the court system and at best left them unable to earn course credit even when their intentions were to remain in school. Not all students interviewed recognized these rigid policies as components of the accountability system, but they were aware that as students, they were having to navigate rules that seemed arbitrary and unhelpful to their progress through school. Few experienced attempts by adults in the school (teachers, counselors, principals, or others) to encourage and make possible their staying in school.

The research reported here is complex, investigating student data, school-level practices, policy analysis, and student experiences, but the findings are simple. We have documented that there is a strong association between high-stakes test-based accountability and large-scale dropping out. This is in large part owing to the system's internal administrative incentives, which reward increased school ratings, even if they are produced at the expense of youth whose test scores are not likely to contribute positively to these production of these indicators.

We have also found that the disaggregation of test scores by ethnic subgroups does not lead to greater equity under a standardized accountability system, as claimed by its advocates. Instead, it marks those students as potential liabilities to the system and to each school's successful ratings under that system. These students come to be seen according to whether they (and their test scores and attendance and behavior records) are assets or liabilities to a school's or district's ratings. The higher the stakes and the longer such a system governs our schools, the more school personnel may come to view students not as children to educate but as potential liabilities or assets for their school's performance indicators and for their own careers.

This study was conducted in Texas, the state whose accountability system became the model for the federal policy governing our nation's schools. The claims of both the state and national system are that they are improving education and that they leave no children behind. Under NCLB, the stakes are even higher because the rewards and sanctions fall to schools and districts in the form of reductions in funding, particularly the reduction of Title I dollars to high-poverty schools. We anticipate that the pressures to view students as assets or liabilities to this system and its production of positive indicators will greatly multiply with this scaling up of sanctions. The triaging of minority youth out of our schools becomes not a side effect of standardized accountability, but an avoidable loss to make the system look successful. It is this shift in thinking about the youth in our schools—as assets or liabilities to the system rather than has having innate worth—that is our more sobering finding. It most directly challenges the rhetoric that such a system is essential for educating all youth, for leaving no children behind.

References

American Educational Research Association. (2000, July) *AERA Position Statement on High-Stakes Testing in Pre-K–12 Education*. Washington, DC: Author.

Anagnostopoulos, D. (2006). "Real students" and "true demotes": Ending social promotion and the moral ordering of urban high schools. *American Educational Research Journal, 43*(1), 5–42.

Arenson, K. W. (2004, May 15). More youths opt for G.E.D. tests, skirting the hurdle of high school. *New York Times*, p. A14.

Balfanz, R. & Legters, N. (2001, January). *How many central city high schools have a severe drop out problem, where are they located, and who attends them?* Paper presented at the "Dropout Research: Accurate Counts and Positive Interventions" Conference, Cambridge MA.

Balfanz, R., & Legters, N. (2004). *Locating the Dropout Crisis*. Baltimore, MD: Center for Social Organization of Schools, John Hopkins University. Retrieved October 21, 2004, from http://www.csos.jhu.edu/tdhs/rsch/Locating_Dropouts.pdf

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis, 24*(4), 305–332.

Carnoy, M., Loeb, S., & Smith, T. (2001). *Do higher state test scores in Texas make for better high school outcomes?* CPRE Research Report No. RR-047. Philadelphia, PA: Consortium for Policy Research in Education.

Clarke, M., Haney, W., & Madaus, G. (2000, January). *High stakes testing and high school completion*. Boston, MA: The National Board on Educational Testing and Public Policy. Retrieved April 24, 2007, from http://www.bc.edu/research/nbetpp/publications/v1n3.html.

Coppola, E. (2007, April 9). *Additive pressures: The multiple effects of the Texas accountability system on high school students' experience in school and their persistence to graduation*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

DeBray, E., Parson, G., & Avila, S. (2003). Internal alignment and external pressure: High school responses in four state contexts. In M. Carnoy, R. Elmore, & L. Siskin (Eds.), *The new accountability: High schools and high-stakes testing* (pp. 55–85). New York: Routledge Falmer.

Denton, D. (2001, January). *Finding alternatives to failure: Can states end social promotion and reduce retention rates?* Atlanta: Southern Regional Education Board. Retrieved February 27, 2006, from http://www.sreb.org/programs/srr/pubs/alternatives/AlternativesToFailure.pdf

Dorn, S. (2003). High-stakes testing and the history of graduation. *Education Policy Analysis Archives, 11*(1). Retreived December 10, 2007 from http://epaa.asu.edu/epaa/v11n1/.

Dorn, S. (2007). *Accountability Frankenstein: Understanding and taming the monster*. Charlotte, NC: Information Age Publishing.

Ekstrom, R.B., Goertz M.E., Pollack, J.M., & Rock, D.A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record, 87*, 356–373.

Fine, M. (1991). *Framing dropouts: Notes on the politics of an urban public high school*. Albany: State University of New York Press.

Fine, M. (2005). High stakes testing and lost opportunities: The New York State Regents Exams. *Encounter, 18*(2), 24–29.

Firestone, W., et al. (2002). The ambiguity of test preparation: A multimethod analysis in one state. *Teachers College Record, 104*(7), 1485–1523.

Gold, B. A. (2007). *Still separate and unequal: Segregation and the future of urban school reform*. New York: Teachers College Press.

Gotbaum, B. (2002, November 21). *Pushing out at-risk students: An analysis of high school discharge figures*. New York: Advocates for Children of New York.

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives, 8*(41) Retrieved from http://epaa.asu.edu/epaa/v8n41/.

Heubert, J.P., & Hauser, R.M. (Eds.). (1999). *High stakes: testing for tracking, promotion, and graduation*. Washington, DC: National Academy Press.

Holmes, C. T. (2006). Low test scores + high retention rates = more dropouts. *Kappa Delta Pi Record, 42*(2), 56–58.

Intercultural Development Research Association (1999). *Missing: Texas youth—dropout and attrition in Texas public high schools*. San Antonio, TX: Author. Retrieved March 20, 2002, from http://www.idra.org/Research/dropout.htm#Dropout.

Jacob, B. A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis, 23*(2), 99–122.

Jacob, B. & Lefgren, L. (2007). *The effects of grade retention on high school completion*. National Bureau of Economic Research Working Paper Series, No. 13514. Retrieved December 11, 2007, from http://www.nber.org/papers/w13514.

Jimerson, S. R., Pletcher, S. M., Graydon, K., Schnurr, B. L., Nickerson, A. B., & Kundert, D. K. (2006). Beyond grade retention and social promotion: Promoting the social and academic competence of students. *Psychology in the Schools, 43*(1), 85–97.

Johnson, R. L. (2007). *Texas public school attrition study, 2006–07: Texas school holding power worse than two decades ago*. San Antonio, TX: Intercultural Development Research Association. Retrieved November 11, 2007, from http://www.idra.org/IDRA_Newsletters/October_2007_School_Holding_Power/Texas_Public_School_Attrition_Study_2006_07/.

Johnson, R., & Onwuegbuzie, A. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26.

Lillard, D. R., & DeCicca, P. P. (2001). Higher standards, more dropouts? Evidence within and across time, *Economics of Education Review, 20*(5), 459–473.

Losen, D., Orfield, G., & Balfanz, R. (2006). *Confronting the graduation rate crisis in Texas*. Los Angeles: The Civil Rights Project. Retrieved July 18, 2007, from http://www.civilrightsproject.ucla.edu/research/dropouts/texas_10-17-06.pdf.

Maxwell, J. (1996). *Using qualitative research to develop causal explanations*. Cambridge, MA: Harvard Project on Schooling and Children.

McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*, 273–298.

McNeil, L. M. (2000). *Contradictions of school reform: Educational costs of standardized testing*. New York: Routledge.

McNeil, L. M. (2001). The harmful effects of the TAAS system of testing in Texas: Beneath the accountability rhetoric. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public schooling* (pp. 127–150). New York: The Century Press.

McNeil, L. M. (2005). Faking equity: High-stakes testing and the education of Latino youth. In A. Valenzuela (Ed.), *Leaving children behind: How "Texas-style" accountability fails Latino youth* (pp. 57–111). New York: State University of New York Press.

McNeil, L. M., & Coppola, E. (2006). Official and unofficial stories: Getting at the impact of policy on educational practice. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp.681–700). Washington, DC: American Educational Research Association.

Miles, M., & Huberman, M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.

Orfield, G. (2004). *Dropouts in America: Confronting the graduation rate crisis*. Cambridge, MA: Harvard Education Press.

Orfield, G., & Ashkinaze, C. (1991). *The closing door: Conservative policy and black opportunity*. Chicago: University of Chicago Press.

Orfield, G. & M. Kornhaber, (2001). *Raising standards or raising barriers? Inequality and high-stakes testing in public education*. New York: The Century Foundation Press.

Orfield, G., Losen, J., Wald, C., & Swanson, B. (2004). *Losing our future: How minority youth are being left behind by the graduation rate crisis*. Cambridge, MA: The Civil Rights Project. Retrieved October 21, 2004, from http://www.civilrightsproject.harvard.edu/research/dropouts/dropouts04.php.

Patton, M. (1990). *Qualitative evaluation and research methods*. (2nd ed.). Thousand Oaks, CA: Sage.

Radigan, J. (2003, April 23). *Working in the margins*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Radigan, J. (2005). *The power of exclusion: Potential high school dropouts speaking out*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Radigan, J. (2007, April 9). *The accountability system's back door: Raising school ratings by losing low-income and minority youth*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Reynolds, A. J. (1992). Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis, 14*(2), 101–121.

Roderick, M., Bryk, A. S., Jacob, B. A., Easton, J. Q., & Allensworth, E. (1999). *Ending social promotion: Results from the first two years*. Charting reform in Chicago, series 1. Chicago: Consortium on Chicago School Research.

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of high school. *American Educational Research Journal, 36*(2), 303–343.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*(3), 583–625.

Rumberger, R. W. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education, 73*(1), 39–67.

Rumberger, R. W. (2004). Why students drop out of school. In G. Orfield (Ed.), *Dropouts in America: Confronting the graduation rate crisis* (pp.131–155). Cambridge, MA: Harvard Education Press.

Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school drop out. *American Journal of Education, 107*, 1–35.

Schemo, D. J. (2003, September 17). Graduation study suggests that some states sharply understate high school dropout rates. *New York Times*, p. B9.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics 6*(2), 461–464.

Sloan, K. (2005). Playing to the logic of the Texas accountability system: How focusing on "ratings"—not children—undermines quality and equity. In A. Valenzuela (Ed.), *Leaving*

children behind: How "Texas-style" accountability fails Latino youth (pp. 153–178). Albany: State University of New York Press.

Smith, F. (1986). *High school admission and the improvement of schooling*. New York: New York City Board of Education.

Texas Code §28.0211 (2007).

Texas Education Agency. (2001). *Grade-level retention in Texas public schools, 1998–99*. Austin, TX: Author.

Texas Education Agency (2002). *Academic Excellence Indicator System 2001–02 state performance report*. Austin, TX: Author. Retrieved October 22, 2003, from http://www.tea.state.tx.us/perfreport/aeis/2002/state.html.

Texas Education Agency. (2003, August). *Secondary school completion and dropouts in Texas public schools 2001–02: Supplemental district data*. Austin, TX: Author. Retrieved March 9, 2007, from http://www.tea.state.tx.us/research/pdfs/0102dist.pdf.

Vasquez Heilig, J. (2006). *Progress and learning of urban minority students in an environment of accountability*. Unpublished doctoral dissertation, Stanford University.

Vasquez Heilig, J. & Darling-Hammond, L. (in press). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Education Evaluation and Policy Analysis*.

Warren, J. R. & Edwards, M. R. (2005). High school exit examinations and high school completion: Evidence from the early 1990s. *Educational Evaluation and Policy Analysis, 27*(1), 53–74.

Wheelock, A. (2003, November). Myopia in Massachusetts: The state's focus on scores harms students and ignores crucial indicators of school quality. *Educational Leadership, 61*(3), 50–54.

Wehlage, G.G., & Rutter, R.A. (1986). Dropping out: How much do schools contribute to the problem? *Teachers College Record, 87*, 374–392.

Young, B. A. (1999). Char*acteristics of the 100 largest public elementary and secondary school districts in the United States: 1997–98* [table]. U.S. Department of Education, National Center for Education Statistics, Washington, DC. Retrieved January 7, 2003, from http://nces.ed.gov/pubs99/1999318/table5.html.

Young, B. A. (2002). *Characteristics of the 100 largest public elementary and secondary school districts in the United States: 2000–01* [table]. U.S. Department of Education, National Center for Education Statistics, Washington, DC. Retrieved January 7, 2003, from http://nces.ed.gov/pubs2002/100_largest/table_01_1.asp.

Appendix A

Descriptive statistics, variables used in school-level regression analyses (1997–2002)

	N	Min.	Max.	Mean	Std. Dev.
School Capacity					
Δ % Fully certified	94	-11.66	14.7	3.15	5.34
Δ % Novice teacher	94	-13.05	4.97	-4.94	3.76
Δ % Teacher turnover	95	-22.37	9.71	-1.47	5.02
Student Progress					
Δ % 9th disappearance	96	-24.26	19.48	2.00	8.86
Δ % 9th retained	96	-14.12	15.53	0.86	5.45
Δ % 9th withdrawal	96	-12.88	23.02	0.73	6.81
Δ % 9th dropout	96	-6.14	4.22	0.05	1.27
School Demographic					
Δ % White	96	-10.27	3.04	-0.31	1.48
Δ % LEP	96	-10.36	12.57	0.11	3.2
Δ % Special education	96	-2.76	6.78	0.60	1.48
Δ % At-risk	96	-12.86	28.05	4.90	7.98

Appendix B

Multinomial logistic regression of TEA accountability changes (1997–2002): coefficients and standard errors

	Model and TEA Rating Change							
	A		B		C		D	
	Falling	Rising	Falling	Rising	Falling	Rising	Falling	Rising
Δ School Capacity								
% Fully Certified	.08 (.05)	.05 (.08)	—	—	—	—	.09 (.07)	.24* (.12)
% Novice Teacher	-.02 (.08)	-.04 (.11)	—	—	—	—	-.16 (.11)	-.08 (.17)
% Teacher Turnover	.05 (.06)	.02 (.08)	—	—	—	—	.10 (.10)	.07 (.13)
Δ 9th Student Progress								
% 9th Disappearance	—	—	-.11 (.06)	.00 (.07)	—	—	-.06 (.08)	.13 (.11)
% 9th Retained	—	—	.22** (.07)	.29** (.09)	—	—	.22* (.10)	.44* (.16)
% 9th Withdrawal	—	—	-.15 (.08)	-.07 (.10)	—	—	-.15 (.10)	.04 (.15)
% 9th Dropout	—	—	-.15 (.24)	-.66* (.32)	—	—	-.44 (.30)	-1.85** (.63)
Δ School Demographics								
% White	—	—	—	—	.20 (.22)	.00 (.30)	.32 (.30)	.68 (.52)
% LEP	—	—	—	—	-.14 (.13)	-.26 (.18)	-.19 (.20)	-.16 (.31)
% Special Education	—	—	—	—	-.16 (.22)	-.64 (.34)	.03 (.29)	-1.20* (.55)
% At-Risk	—	—	—	—	.20** (.07)	.04 (.09)	.19* (.09)	-.04 (.14)

Numbers in parentheses are standard errors. *$p < .05$ **$p < .01$

About the Author

Linda McSpadden McNeil
Rice University

Eileen Coppola
Rice University

Judy Radigan
Rice University

Julian Vasquez Heilig
University of Texas at Austin

Email: lmcneil@rice.edu

Linda McSpadden McNeil is professor of education and director of the Rice University Center for Education. She is the author of Contradictions of School Reform: Educational Costs of Standardized Testing (Routledge, 2000). She holds a doctorate in Curriculum and Instruction from the University of Wisconsin-Madison. Her research encompasses educational quality and equity, and the education of urban youth in the context of standardization and globalization.

Eileen Coppola is a Research Scientist at the Rice University Center for Education and a lecturer in Education at Rice. She is the author of Powering Up: Learning to Teach Well with Technology (Teachers College Press, 2004). Her doctorate from the Harvard University Graduate School of Education is in Administration, Policy, and Social Planning in the Urban Superintendents Program. Her recent research is on the effects of policy on classrooms and student experience, and on the quality of students' engagement with schooling.

Judy Radigan is a Research Scientist at the Rice University Center for Education and a lecturer in Education at Rice. She holds a PhD from the University of Houston in Educational Psychology and has extensive experience as a teacher and principal in urban schools. She is devoted to researching the lives of youth in schools and supports equitable educational opportunities.

Julian Vasquez Heilig is an Assistant Professor of Educational Policy and Planning at the University of Texas at Austin. He obtained his Ph.D. in Educational Administration and Policy Analysis from Stanford University. He holds masters degrees from Stanford and the University of Michigan. His current research includes quantitatively and qualitatively examining how high-stakes testing and accountability-based reforms and incentive systems impact minority students.

EDUCATION POLICY ANALYSIS ARCHIVES http://epaa.asu.edu

EDUCATION POLICY ANALYSIS ARCHIVES **http://epaa.asu.edu**