

## Holding Back Students Damages Their Educational Progress: An Advocacy Report

A review of the research on the efficacy of retention—the blanket strategy of using a single test to determine if a student should be held back—demonstrates that retention impedes the educational progress of children and leads primarily to lower achievement and higher dropout rates. Indeed, few issues have generated such an overwhelming consensus among professionals as on the negative effects of retention.

Nevertheless, under political pressure for a measure that at least appears to address long-standing educational problems, in March 2004 New York City's Mayor Michael Bloomberg forced through a policy of retention based on a single measure of performance for third graders. This was done despite widespread protest from advocates across the city and opposition from a majority of members of the Panel on Education Policy. The policy was passed only because three panel members were removed hours before the vote. In September 2004 the same policy passed for fifth graders.

This latest fight to stop a blanket retention policy based on standardized tests failed here in New York City. Still, our hope is that by describing both the research that demonstrates why the use of a single test to determine whether a child is held back is so damaging to students, particularly for students of color, and the advocacy tools used to oppose such a policy, we can offer to advocates in other cities and school districts arguments effectively to block such detrimental policies.

### The Folly of a Single Test to Determine Holdover

One of the inherent problems with the third-grade and fifth-grade retention policy is the use of a single test to determine if a child will be held back. Test makers and educational researchers conclude that a single test should not be the basis for significant educational decisions. The American Educational Research

Association, the nation's largest professional organization devoted to the scientific study of education, opposes their use in this way, as do the National Board on Educational Testing, the International Reading Association, and the National Council of Teachers of Mathematics, which argues that “far-reaching and critical educational decisions should be made only on the basis of multiple measures.”<sup>1</sup> The Standards for Educational and Psychological Testing, developed by the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education, contain the following statement: “Any decision about a student's continued education, such as retention, tracking, or graduation, should not be based on the results of a single test, but should include other relevant and valid information.”<sup>2</sup>

The National Academy of Sciences published a comprehensive report explaining in detail why the use of high-stakes testing to determine the next step in a student's education is intellectually indefensible as well as counterproductive.<sup>3</sup> As the authors point out, “A student's score can be expected to vary across different versions of a test . . . as a function of the particular sample of questions asked and/or transitory factors, such as the student's health on the day of the test. Thus, no single test score can be considered a definitive measure of a student's knowledge.”<sup>4</sup>

Harcourt and CTB McGraw Hill, the two largest standardized test producers, and the developers of New York City's third-grade reading and mathematics examinations, oppose the use of their tests as the exclusive criterion for decisions about retention because they can never be a reliable or complete measure of what students may or may not know. Harcourt writes that achievement test scores “should be just one of many factors considered and probably should receive less weight than factors such as teacher obser-

<sup>1</sup>American Educational Research Association, Position Statement Concerning High-Stakes Testing in Pre-K–12 Education (2000), [www.aera.net/policyandprograms/?id=478](http://www.aera.net/policyandprograms/?id=478); International Reading Association, Summary of Position Statement from High-Stakes Assessments in Reading (1999), [www.reading.org/downloads/positions/ps1035\\_high\\_stakes.pdf](http://www.reading.org/downloads/positions/ps1035_high_stakes.pdf); National Council of Teachers of Mathematics, Position Statement on High-Stakes Testing, [www.nctm.org/about/position\\_statements/highstakes.htm](http://www.nctm.org/about/position_statements/highstakes.htm).

<sup>2</sup>American Educational Research Association, *supra* note 1.

<sup>3</sup>COMMITTEE ON APPROPRIATE TEST USE, NATIONAL RESEARCH COUNCIL HIGH STAKES: TESTING FOR TRACKING, PROMOTION, AND GRADUATION (Jay P. Heubert & Robert M. Hauser eds., 1999), available at [www.nap.edu/catalog/6336.html](http://www.nap.edu/catalog/6336.html).

<sup>4</sup>*Id.* at 3.

vation, day-to-day classroom performance, maturity level, and attitude.”<sup>5</sup>

As with all standardized tests, a substantial margin of error is inescapable because of the nature of these examinations. Thus a number of students who would pass if the statistical uncertainties involved were taken into account are likely to fail.<sup>6</sup> The tests themselves may be flawed or may be scored incorrectly, as has occurred in the recent past. In such cases, more students would be unfairly held back, and their futures put at unnecessary risk.

New York City is not immune to these potential errors. During the April 2004 third-grade examination, which was the new sole measure of promotion, officials reported that 1,300 students had previously seen test questions because they were used on earlier administrations of the test. This cast doubt on the validity of those test scores.<sup>7</sup>

### Historically Ineffective Retention Policies Based on a Single Test

A retention policy based on a single test has been tried before in New York City and a number of other school districts. Research data on the New York City program demonstrated that it was a clear failure, and the program was ended, but not until after huge sums of money were expended. Ten years after Chancellor Frank J. Macchiarola launched the Promotional Gates program in 1981, it was eliminated. Under Macchiarola’s plan 25,000 fourth through seventh graders were held back the first year. Low scores on citywide reading examinations were the basis for the retention decision; mathematics scores were included later.

The plan brought with it a huge financial burden. Promotional Gates required the hiring of an additional 1,100 teachers for the newly created retention classes capped at 18 students, and a summer school program. The additional teachers hired alone cost between \$40 million

and \$70 million, yet student outcomes were negative. The average summer school student made no improvements on their test score performance. After two years, retained students still showed no significant improvements over low-achieving counterparts who were promoted. Moreover, long-term follow-up showed that 40 percent of the students who were retained eventually dropped out, compared to 25 percent of those with similar test scores who had been promoted. According to Ernest House, one of the authors of an evaluation mandated by the mayor’s office, “the Promotional Gates Program had retained tens of thousands of students at huge dollar and human costs without benefits.”<sup>8</sup>

In light of this failure, on September 11, 1991, the Board of Education adopted a resolution eliminating the Gates program. Recognizing the failure of the philosophy behind Promotional Gates, the board discontinued the program because “it did not sufficiently improve the achievement levels of participating grade 4 and 7 students.”<sup>9</sup> The 1991 resolution conceded that the Promotional Gates program had made them more likely to drop out.<sup>10</sup>

The dismal results of the Promotional Gates program was unfortunately replicated in the single-test policy enacted in the Chicago public schools, which also failed.<sup>11</sup> In 1996, the Chicago public school system instituted a policy of promoting students on the basis of their performance on the Iowa Tests of Basic Skills. The plan went into effect for thousands of third, sixth, and eighth graders, and thousands of students were retained.

A recent report by the Consortium on Chicago School Research tracked the effects of the policy on these students retained as a result. Their findings reinforce the overwhelming professional research consensus that single-test promotional standards do not work and instead have deleterious and discouraging effects. The researchers essentially found that in the third grade the promotion policy had no effect on student performance.

<sup>5</sup>HARCOURT BRACE EDUCATIONAL MEASUREMENT, *STANFORD ACHIEVEMENT TEST SERIES: GUIDE FOR ORGANIZATIONAL PLANNING* 43–44 (9th ed. 1997).

<sup>6</sup>E.g., an analysis showed that those students who really belonged at the fiftieth percentile of the widely-used Stanford 9 test would be expected to score within five points of that mark only about 30 percent of the time in mathematics and only 42 percent of the time in reading. David Rogosa, *How Accurate Are the STAR National Percentile Rank Scores for Individual Students?—An Interpretive Guide* (1999), [www-stat.stanford.edu/~rag/ed351/drrguide.pdf](http://www-stat.stanford.edu/~rag/ed351/drrguide.pdf).

<sup>7</sup>David M. Herszenhorn, *Retest Is Option for 3rd Graders Who Got Peek*, *NEW YORK TIMES*, April 29, 2004, at B5.

<sup>8</sup>Ernest R. House et al., *An Audit of the Evaluation of New York City’s Promotional Gates Program* (four reports: Oct. 1981, Feb. 1982, April 1982, Oct. 1982). House summarizes the report’s conclusions in *The Predictable Failure of Chicago’s Student Retention Program* (1998), available at [www.designsforchange.org/pdfs/houseChicago.pdf](http://www.designsforchange.org/pdfs/houseChicago.pdf). See also the negative evaluation of the Promotional Gates program by R.D. Gampert & P. Opperman, *Longitudinal Study of the 1982–83 Promotional Gates Students* (1988), cited in *COMMITTEE ON APPROPRIATE TEST USE*, *supra* note 3, at 128 n.13.

<sup>9</sup>New York City Board of Education Resolution, Sept. 11, 1991, at 1.

<sup>10</sup>*Id.* at 2. See also Joseph Berger, *Fernandez to End a Policy on Holding Pupils Back*, *NEW YORK TIMES*, Aug. 3, 1990, at A1.

<sup>11</sup>ERNEST R. HOUSE, *THE PREDICTABLE FAILURE OF CHICAGO’S STUDENT RETENTION PROGRAM* 17 (1998). For findings of recent reports tracking the effects of the policy on students, see MELISSA RODERICK & JENNY NAGOAKA, *ENDING SOCIAL PROMOTION: THE EFFECTS OF RETENTION 41* (Consortium on Chicago School Research, *Charting Reform in Chicago Series*, 2004).

Sixth graders experienced different, much worse, outcomes. The retained students in the sixth-grade cohort performed worse than those promoted. In the first post-Gates program year, retained sixth-grade students achieved only a 0.82 unit growth in their reading abilities, while the similarly low-achieving promoted students had a growth of 1.19 units—a significant 0.37 difference. Two years after the gate, again there was an even greater disparity of growth—a 0.44 difference. Retained sixth graders performed almost 25 percent worse than low-achieving, similarly performing sixth-grade students who were promoted.<sup>12</sup> The consortium also found that the performance of the lowest achieving students, who were not part of any cohort, deteriorated after being retained. While this does not prove that retention caused that deterioration, it leads to the reasonable conclusion that retention did not help or stop the decline in performance.

### Retention Policies' Disproportionate Effect on Black and Hispanic Students and Increased Dropout Rates

The large-scale retention policy is not only counterproductive and extremely expensive but also inherently inequitable. The practice of retaining large numbers of students on the basis of test scores alone is likely to have a disproportionate effect on those who are poor and minority students. Moreover, the policy of using high-stakes tests to make retention decisions has been shown to be more commonly used in school districts with high percentages of black and Hispanic students compared to the rest of the nation. Since research shows these policies on balance to be harmful to students who are subjected to them, their use appears to exacerbate rather than ameliorate racial and class differences.

To give some concrete examples in New York City, District 12, with a minority population of over 93 percent had only 19 percent of students achieve a level 3 or 4 on the 2004 mathematics examination. Conversely District 2 in Manhattan, which has a 70 percent white population, had over 60 percent of its students score 3

or 4 on the same examination.<sup>13</sup> Figures such as these suggest that more minority than white students are affected by promotional policies.

Besides exacerbating racial and class differences, retention leads to increased dropout rates, as was seen during the Promotional Gates program. Numerous studies on retention conclude that students who are held back are much more likely to drop out eventually. One study found that a student retained once was 40 percent to 50 percent more likely to drop out of school, and 90 percent more likely to drop out if retained twice.<sup>14</sup> After controlling for student background and academic achievement, a longitudinal study of more than 12,000 students found that, being held back before the eighth grade increased the likelihood of dropping out by the twelfth grade by more than 200 percent. Furthermore, “students who were held back before the 8th grade were more than four times as likely as students who were not held back to not complete high school or receive a GED” six years later.<sup>15</sup>

### Protest Against Retention

With such a strong body of evidence demonstrating that blanket retention policies do not work, fighting such policies would seem unnecessary: why would educators adopt a policy that has failed time and again, and always at great expense? However, education policy is often intertwined with the politics of the school district. In New York City that has certainly always been the case, and with clear mayoral control granted to New York City's mayor as of June 2002, the issue of retention arose again as a possible “quick fix” to longstanding problems of lagging student achievement.<sup>16</sup> Recognizing the need for action to prevent this new blanket retention policy proposed in 2004, advocates used a number of strategies effectively. These included sign-on letters, use of the media for outreach and to influence public opinion, and petitions.

As described above, the Mayor's retention policy was attempted previously as the Promotional Gates pro-

<sup>12</sup>RODERICK & NAGOAKA, *supra* note 11.

<sup>13</sup>NEW YORK CITY DEPARTMENT OF EDUCATION, GRADE 5 MATH BY DISTRICT: 1999–2004, [www.nycenet.edu/daa/2004Math3567/excel/MATH%205%20by%20District.xls?id\\_](http://www.nycenet.edu/daa/2004Math3567/excel/MATH%205%20by%20District.xls?id_), PERFORMANCE ON GRADE 5 ELA BY DISTRICT: 1999–2004, [www.nycenet.edu/daa/2004ela3567/excel/ELA%205%20by%20District.xls](http://www.nycenet.edu/daa/2004ela3567/excel/ELA%205%20by%20District.xls); NEW YORK CITY DIVISION OF PLANNING, NEW YORK CITY PUBLIC SCHOOLS ENROLLMENT AND DEMOGRAPHIC TRENDS: 1990–2002, [www.nyc.gov/html/dcp/pdf/pub/schilbronx.pdf](http://www.nyc.gov/html/dcp/pdf/pub/schilbronx.pdf), [www.nyc.gov/html/dcp/pdf/pub/schmanhattan.pdf](http://www.nyc.gov/html/dcp/pdf/pub/schmanhattan.pdf).

<sup>14</sup>D. Mann, *Can We Help Dropouts? Thinking About the Undoable*, in *SCHOOL DROPOUTS: PATTERNS AND POLICIES 3–19* (G. Natriello ed., 1987). Shane R. Jimerson et al., *Winning the Battle and Losing the War: Examining the Relation Between Grade Retention and Dropping out of High School*, 39 *PSYCHOLOGY IN THE SCHOOLS* 12 (2002).

<sup>15</sup>Russell W. Rumberger & Katherine A. Larson, *Student Mobility and the Increased Risk of High School Dropout*, 107 *AMERICAN JOURNAL OF EDUCATION* 1 (1998).

<sup>16</sup>In 2002 only 37.6 percent of English-proficient students were performing at levels of proficiency on the state and city mathematics examinations, and less than 10 percent of English language learners scored at levels of proficiency. Similar results were found on the English language arts examination, where 41.5 percent of English-proficient students were performing at levels of proficiency and only 5.8 percent of English language learners were scoring at levels of proficiency. New York City Department of Education, Report on the 2003 Results of the State English Language Arts (ELA) Tests, and the City ELA and Math Assessments (2003), [www.nycenet.edu/daa/2003reading/2003\\_Test\\_Results.pdf](http://www.nycenet.edu/daa/2003reading/2003_Test_Results.pdf).

gram and was officially discontinued in 1991 by the New York City Board of Education. Only eight years after this last full-scale retention policy in New York City was proved to be a dismal failure, which cost hundreds of millions of dollars, a new retention policy was proposed by then Chancellor Rudy Crew. An outcry from the advocacy community ensued. This included litigation filed by Advocates for Children of New York.<sup>17</sup> The litigation did not itself stop the retention policy, but the overall outcry made the Chancellor quickly roll back his original proposal for a blanket retention policy. Instead he proposed a three-pronged policy for determining promotion: standardized test scores, school work, and attendance levels. If a student received a score of between 2 and 4 on state and city tests, had class work that was on grade level, and was attending school 90 percent or more of the time, the student moved on to the next grade.<sup>18</sup> In fact, according to the then revised chancellor's regulation, if the student met two out of the three criteria, the student was allowed to move forward. This policy remained unchanged until early 2004.

As direct mayoral control of the public schools had been established in New York City, the mayor staked his mayoralty on their improvement. Although he had mentioned the issue of "ending social promotion" in some of his speeches, no concrete policy emerged until January 2004. At this point Mayor Bloomberg outlined a blanket retention policy based on the outcome of one test. The delivery of additional academic supports to struggling students was added to the proposal.<sup>19</sup> A newly reconfigured board of education, renamed the Panel on Educational Policy and composed of a majority of mayoral appointees, was slated to approve or reject the policy in early March 2004.

Advocates marshaled their resources to oppose this policy, which was clearly detrimental to most of the city's students. Advocates for Children and Class Size Matters released a letter—signed by over 100 academ-

ics, heads of organizations, and experts on testing from throughout the nation—in opposition to the mayor's announced policy to hold back third-graders on the sole basis of their scores on standardized tests.<sup>20</sup>

At a press conference, parents of at-risk third graders spoke eloquently against the proposed policy and put further pressure on the administration. Before the panel's vote, the National Center for Schools and Communities, along with New York University's Institute for Education and Social Policy, came out with a report again demonstrating both the lack of effectiveness of blanket retention policies and their high cost.<sup>21</sup> Members of the panel who opposed the policy disseminated the widespread research to their fellow members. The day before the vote clearly a majority of members would vote against the mayor's policy. The advocacy campaign, which presented the results of over twenty-five years of research data, appeared to have worked. However, hours before the vote was to take place the mayor removed two of his appointed members. Concurrently the Staten Island borough president, apparently at the mayor's behest, removed a third. All three had been planning to vote against the proposal. The three newly appointed members voted unsurprisingly for the retention proposal, and so it passed on March 15, 2004.

However, measures taken by organizations opposed to this policy were not completely in vain. As a result of collective action, an appeals process requiring the review of class work and teacher recommendations of all children slated for holdback because of their test scores under this new policy was added to the policy and raised at least some hope of promoting students under the "gate" set by the administration.<sup>22</sup> On a region-by-region basis, appeals were granted to between 25 percent and 60 percent of third graders.

In August 2004 the mayor and chancellor introduced the same policy proposal for fifth graders. With only

<sup>17</sup>*H.P. v. Crew*, No.118294/1999 (N.Y. Sup. Ct., N.Y. County, Sept. 25, 2000).

<sup>18</sup>The New York State English language arts and mathematics examinations are scored on a scale of 1-4, with 1 the lowest score and 4 the highest score. According to New York State law, 3 (proficient) and 4 (advanced) are the only passing scores. New York City considers students achieving at 2 (basic) as passing and ready to move onto the next grade.

<sup>19</sup>New York City Panel on Education Policy Resolution, March 15, 2004, at 1.

<sup>20</sup>Letter from Advocates for Children and Class Size Matters, to Mayor Michael Bloomberg and Chancellor Joel I. Klein (Feb. 11, 2004), [www.classsizematters.org/retentionletter.html](http://www.classsizematters.org/retentionletter.html). Signers include four past presidents of the American Education Research Association, the nation's premier organization of educational researchers; three members and the study director of the National Academy of Sciences Committee on the Appropriate Use of Educational Testing; and two members of the Board on Testing and Assessment of the National Research Council.

<sup>21</sup>INSTITUTE FOR EDUCATION AND SOCIAL POLICY, STEINHARDT SCHOOL OF EDUCATION, NEW YORK UNIVERSITY, AND NATIONAL CENTER FOR SCHOOLS AND COMMUNITIES, FORDHAM UNIVERSITY, FIRST, DO NO HARM: A RESPONSE TO THE PROPOSED NEW YORK CITY THIRD GRADE RETENTION POLICY (2004).

<sup>22</sup>Elissa Gootman, *Appeals Process for 3rd Graders Explained*, NEW YORK TIMES, Feb. 27, 2004, at B6.

three weeks to respond, and a panel vetted to vote for such a policy, Advocates for Children still worked to oppose this new retention policy for fifth graders. A new letter opposing the policy and signed by over twenty-five academics and national experts on testing and advocates across New York City was sent to the mayor and panel members. The vote however was a foregone conclusion. The measure overwhelmingly passed on September 27, 2004. This time the measure provided funds for evaluating the success of the program; this evaluation will be closely tracked by Advocates for Children.



With such overwhelming evidence of the negative effects of retention, why school administrators would want to implement a program of retention based on a single test is puzzling from an educational policy perspective. The program failed to help New York City students in past and has failed to assist students in school districts across the country for the last twenty-five years.

The single-test retention policy is not supported by the testing companies or by the research and disproportionately affects black and Hispanic students. The funds spent on retaining students would be better used to provide programs that have been demonstrated to improve student performance; they might offer smaller classes, increased access to prekindergarten, and intensive intervention for students who have fallen or are at risk of falling behind. Other school districts should heed the evidence and use their funds more wisely and for the betterment of their at-risk students.

**Jill Chaifetz**

*Executive Director*

Advocates for Children of New York  
151 W. 30th St., 5th Floor, New York, NY 10001  
212.947.9799; jchaifetz@advocates forchildren.org

**Rachel Kravitz**

*Americorps Vista Member/Children's  
Program Associate*

Advocates for Children of New York  
rkravitz@advocates forchildren.org