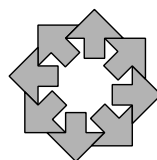


This is a new format for presenting findings from the Consortium on Chicago School Research, departing from our regular, more comprehensive, in-depth studies. As the name suggests, this is a short report focusing on a single topic. The data brief is designed to provide new data on a particular issue, in a timely fashion. Because data briefs are not comprehensive studies, like our various report series, we limit our discussion of findings to summarizing the key results.



THE CONSORTIUM ON
CHICAGO SCHOOL
RESEARCH

Research Data Brief

Adjusting Citywide ITBS Scores for Student Retention in Grades Three, Six, and Eight

December 1998

Research Question

When the Chicago Public Schools released the Iowa Tests of Basic Skills (ITBS) achievement test scores in the summer of 1998, many commentators suggested that the scores had been inflated by the Reform Board's new promotion and retention policy (see *Catalyst*, June 1998, for example). Critics argued that the policy artificially raises test scores because retained students, who would have scored poorly at the higher grade if they had been promoted, instead are retaking the test at the lower grade. In fact, researchers studying retention elsewhere have noted such findings.¹

Our purpose is to illuminate that issue. We are asking how the 1998 ITBS scores of retained students in third grade (15 percent of the students tested), sixth grade (10 percent), and eighth grade (4.6 percent) affect the citywide results in those grades and in the subsequent grades.²

Interpreting school achievement scores and score reports becomes more difficult because of the change in promotion and retention practices. The policy changes the composition of students within grade levels, so that this year's group of third graders is very different from the group of third graders the year before the policy began. Because the groups of students have become so different, it becomes harder to know whether scores are really up or down. For example, are the fourth grade

John Q. Easton
Brian Jacob
Stuart Luppescu
Melissa Roderick

scores up in 1998 just because the lowest scoring students are repeating third grade rather than being included in the fourth grade?

The new CPS promotion and retention policy raises a number of important research questions, many of which have been extensively described in a large body of research literature. Consortium researchers, including Melissa Roderick and John Easton, have recently begun a three-year study to investigate these questions and will report their results on a regular basis beginning in the winter of 1999. The limited data reported here are specifically aimed at helping understand how significant numbers of retained students in grades three, six, and eight affect the citywide results in those grades and in the subsequent grades.

Findings

In third, sixth, and eighth grades, when we remove the scores of retained students, the 1998 test scores are higher than reported in both reading and math. In fourth and seventh grades, when we include the retained students as if they had taken the test in those grades, the scores are lower than reported. Compared with the 1997 scores, the adjusted third, sixth, and eighth grade scores are up, while the fourth and seventh grade scores, with the exception of seventh grade math, are not.

But even after these adjustments, the overall systemwide scores are still up from 1997. Did the promotion/retention policy seriously distort the 1998 ITBS test score results? In our judgment, it did not.

Overall, the 1998 reading scores for the CPS system remain about the same after being adjusted for the effects of retained students (see Figure 15 on page 8). Before the adjustment, 34.3 percent of students scored at or above national norms in reading; after the adjustment, the number is 34.4 percent. Both of these numbers are higher than 1997, when 30.3 percent of students were at or above national norms.

In math, the 1998 unadjusted score of 39.3 percent falls slightly to 38.7 percent after the adjustment. But the adjusted score still represents an increase from the 1997 score of 35.9 percent.

This study also shows that grouping students by age instead of grade is a viable alternative for tracking their scores. The age groupings mirror the grade group findings, with the greatest improvements among students most directly affected by the policy. This pro-

cedure will be useful in the future as the composition of grades becomes increasingly complex.

Research Methods

This data brief examines Iowa Tests of Basic Skills score results in the Chicago Public Schools in light of recent changes in the student promotion policy. Since fall 1996, students at selected grade levels have been required to attain a specified score on the ITBS in both reading comprehension and mathematics. Students who do not reach the criterion in the spring testing program are required to attend an intensive summer school program. At the end of summer school, students are re-tested and subsequently promoted if they meet the criterion and retained in the same grade if they do not. This new promotion and retention policy was fully implemented in the spring, summer, and fall of 1997 for third, sixth, and eighth grade students. The spring 1998 test scores, therefore, were the first to be affected significantly by new patterns of retention and promotion. High school students were also affected, though this data brief does not discuss those students.

This study re-analyzes ITBS data provided by the Chicago Public Schools Office of Accountability. We have adjusted the spring 1998 data for effects of retained students. In addition, we offer an alternative score reporting method based on students' age that may be informative in the future. In all of the analyses, we follow the same rules as CPS for including students in the reporting.

Adjustment Technique: Separating Retained Students from Non-retained Students in Grades Three, Six, and Eight. In this procedure, we identified third, sixth, and eighth grade students in the 1998 test file who were repeating those grades. These are students who were retained in the fall of 1997, the first year of full implementation of the policy. We then separated the third, sixth, and eighth grade scores into two groups: students in the respective grade for the first time, and students in the grade for the second time (that is, the retained students). We removed the retained students' scores from grades three, six, and eight.

The second step in this adjustment involved estimating scores for the retained third and sixth grade students on the fourth and seventh grade test levels. Without the promotion policy, these students would have taken the

test in the higher grade, so we wanted to estimate what their score would have been in that grade. In order to do this, we used our previous equating study that linked different levels and forms of the ITBS. (This study is reported in detail in *Academic Productivity of Chicago Public Elementary Schools*, March 1998.³) In the simplest terms, this process entailed making empirically based statistical predictions of what students who took the third grade test would have scored if they had taken the fourth grade test, and similarly for sixth graders and the seventh grade test. (Unfortunately, we are unable to move from the ITBS to the TAP, the high school achievement test, so we could not adjust ninth grade scores for retained eighth graders.) We then recalculated the fourth and seventh grade scores including the predicted scores of retained third- and sixth-grade students.⁴ Note that we made this adjustment only in 1998, the first year that significant numbers of third, sixth, and eighth grade students were held back under the new promotion policy.

Alternative Reporting Technique: Reporting Scores by Age Group, Rather Than by Grade. In addition to the statistical adjustment strategy described above, we report score trends by age groups as well as by grades. As noted above, the changing promotion policy results in changes in the composition of students within grades over time, therefore making annual grade-to-grade comparisons more difficult. Other factors can also make the grade-to-grade comparisons complex. For example, the required age for eligibility to enter kindergarten changed in the early 1990s. This too can affect scores within grades and confound the annual comparisons. The age grouping also controls for retention that may have occurred in grades other than three, six, and eight.

Other testing programs, including the National Assessment of Educational Progress, control for these changing factors by reporting student scores in age groups rather than by grade. Even though it is still a cross-sectional comparison, age groups at different points in time can be more similar to each other than grade groups are.

To define the age groups, we began on April 15, which is about when most of the citywide testing has occurred over the years. All students whose birthdays are within six months on either side of April 15 are considered part of a specific age group. For example, students who turned 10 years old in the year where April 15 is the midpoint are included in the 10-year-old group.

Use of the Mean as the Statistical Criterion. Because of our previous research on the topic (see *Academic Productivity*), we have elected to report the mean grade equivalent in Figures 1-12, rather than either the median grade equivalent or the percent of students scoring at or above the national norms. Though both of these other statistics can provide useful information about the school system, the mean (arithmetic average) is a superior statistic in examining trends and other changes over time because it takes into account the performance of all students, whereas the other statistics are only sensitive to the performance of selected students. Students right at the middle of the distribution have the greatest influence on the median, and students closest to the national average likewise affect the percent at or above grade level. The mean, on the other hand, is sensitive to the scores of all students. Though we are using the mean in this report, we have also calculated our results in terms of the percent of students at or above grade level and display these results in Figure 15. These results, which are most directly comparable to CPS reports, are similar to the results using the mean grade equivalent.

Effects of Adjustment on Score Trends

The results of our analyses are presented graphically in Figures 1 to 12. Accompanying each graph is a brief interpretive statement.

Following the graphs are complete data tables. Figure 13 contains the average grade equivalent ITBS score by grade for 1990 to 1998, with adjustments made in the grades most directly affected by the new promotion policy in 1998.

Figure 14 contains the average ITBS scores by age group. The age groupings control for changes in retention patterns at all grade levels and in all years, not just for grades three, four, six, seven, and eight in 1998 as in Figure 13.

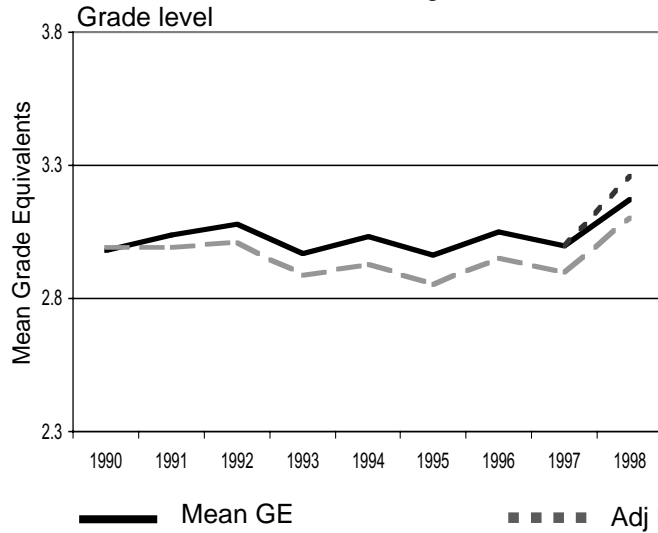
Figure 15 contains the percent of students scoring at or above grade level on the ITBS. This table is most directly comparable to the CPS reports.

See the Consortium on Chicago School Research website (<http://www.consortium-chicago.org>) for additional information, including number of students tested and average scores for retained students.

Third Grade/Nine Year Olds ITBS Mean Scores

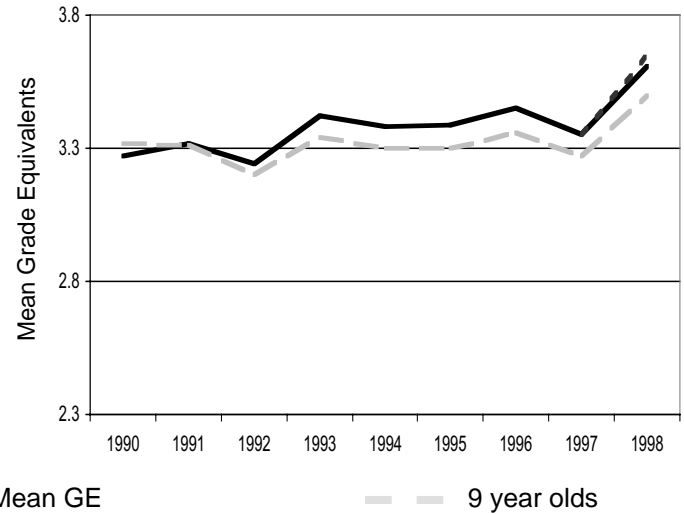
Figure 1

Reading



Math

Figure 2

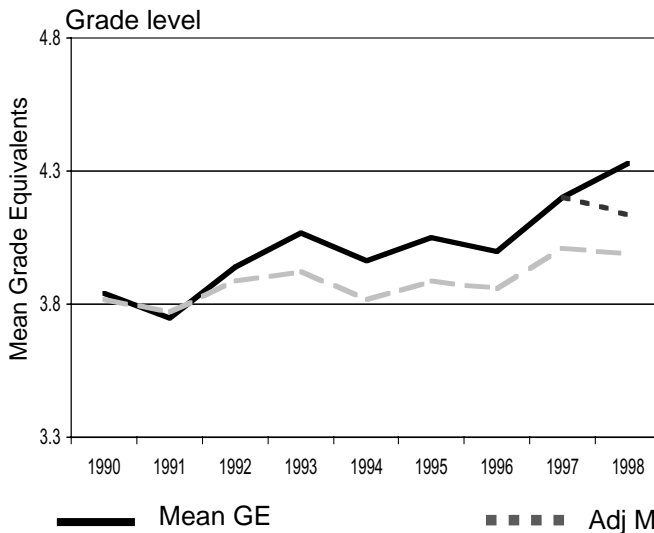


Grade Three/9-Year-Old Students. In both reading and math, the adjusted scores are higher than the unadjusted scores. Removing the retained students increases the scores by approximately one month (in terms of mean grade equivalent) in reading and one-half month in math. Note that the 9-year-old age group trend is nearly parallel to the adjusted third-grade trend. (The age group scores are lower than the grade scores because they include students from lower grades, who will tend to score lower.) Adjusted 1998 scores are higher than 1997 scores in both reading and math.

Fourth Grade/Ten Year Olds ITBS Mean Scores

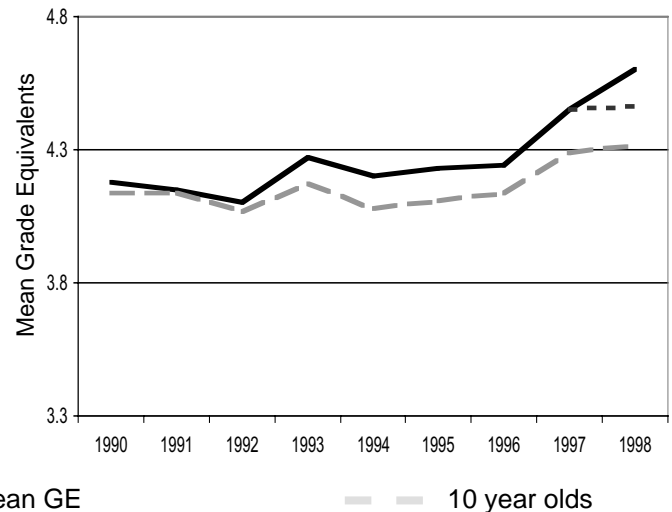
Figure 3

Reading



Math

Figure 4

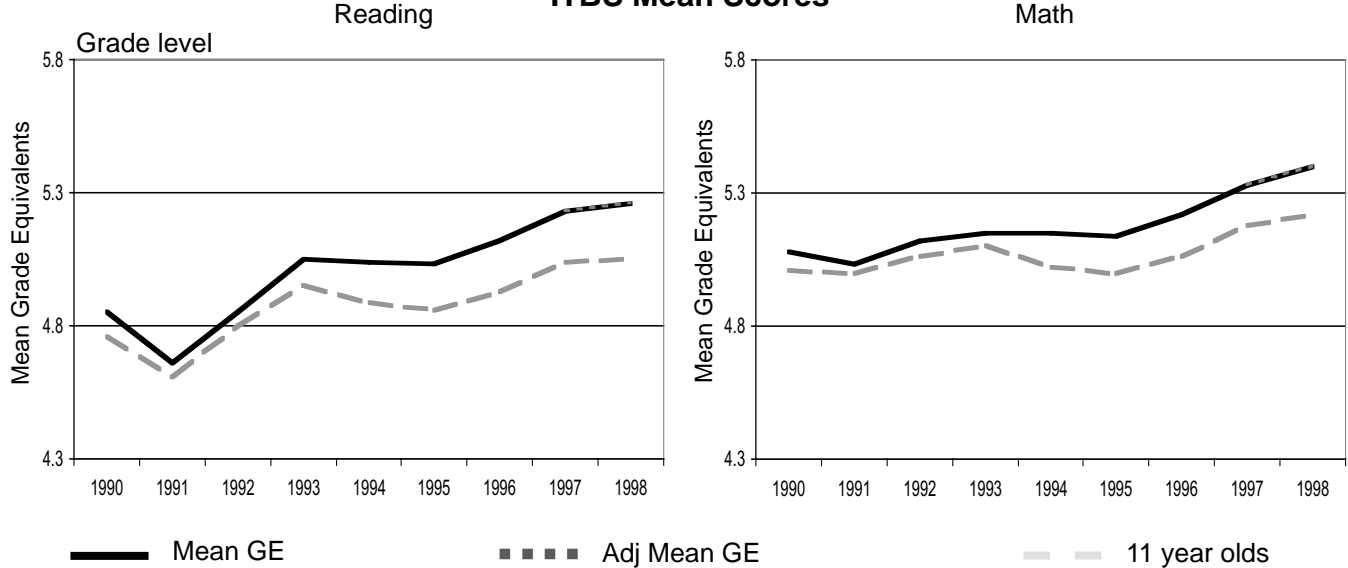


Grade Four/10-Year-Old Students. The adjustments in fourth grade scores work in the opposite direction. Adjusted 1998 scores are lower than the unadjusted scores by about two months in reading and one and one-half months in math, since the weaker retained students are added in here. Again, the age group scores are nearly parallel to the adjusted grade group. Adjusted 1998 scores are lower than 1997 scores in reading and about the same in math.

Fifth Grade/Eleven Year Olds ITBS Mean Scores

Figure 5

Figure 6

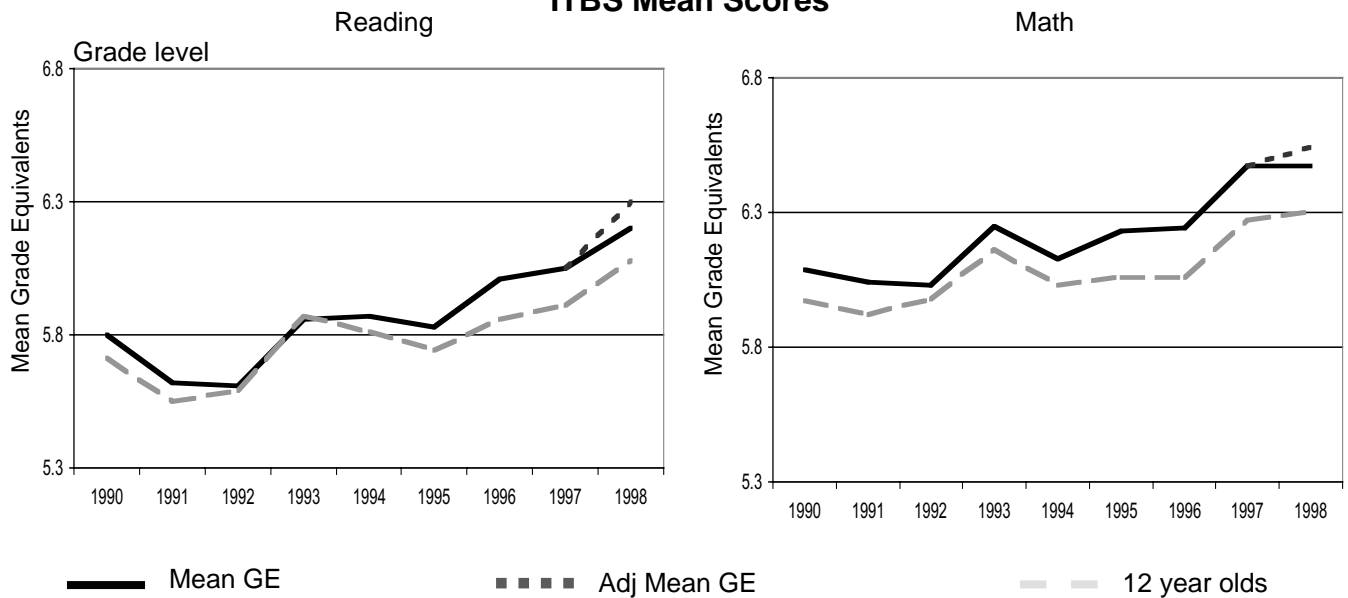


Grade Five/11-Year-Old Students. Fifth grade is the only elementary grade where 1998 scores are completely unaffected by our adjustment for retained students. We did not adjust for them because CPS policy focused on grades three, six, and eight. Fifth grade scores will be affected in 1999 when students retained in the fall of 1997 will not be included in that year's fifth grade testing.

Sixth Grade/Twelve Year Olds ITBS Mean Scores

Figure 7

Figure 8

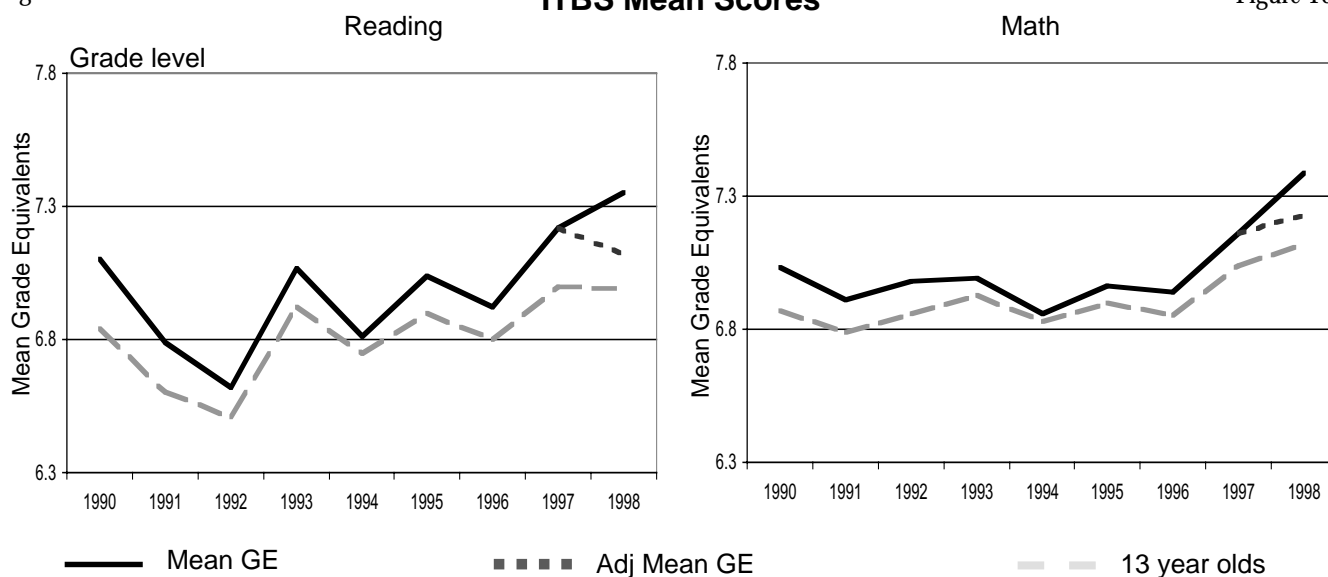


Grade Six/12-Year-Old Students. We see the same pattern here as third grade: the adjustment brings scores up, though not quite as much, and again the adjustment is greater in reading than in math. Also, adjusted 1998 scores are higher than 1997 scores.

Seventh Grade/Thirteen Year Olds ITBS Mean Scores

Figure 9

Figure 10

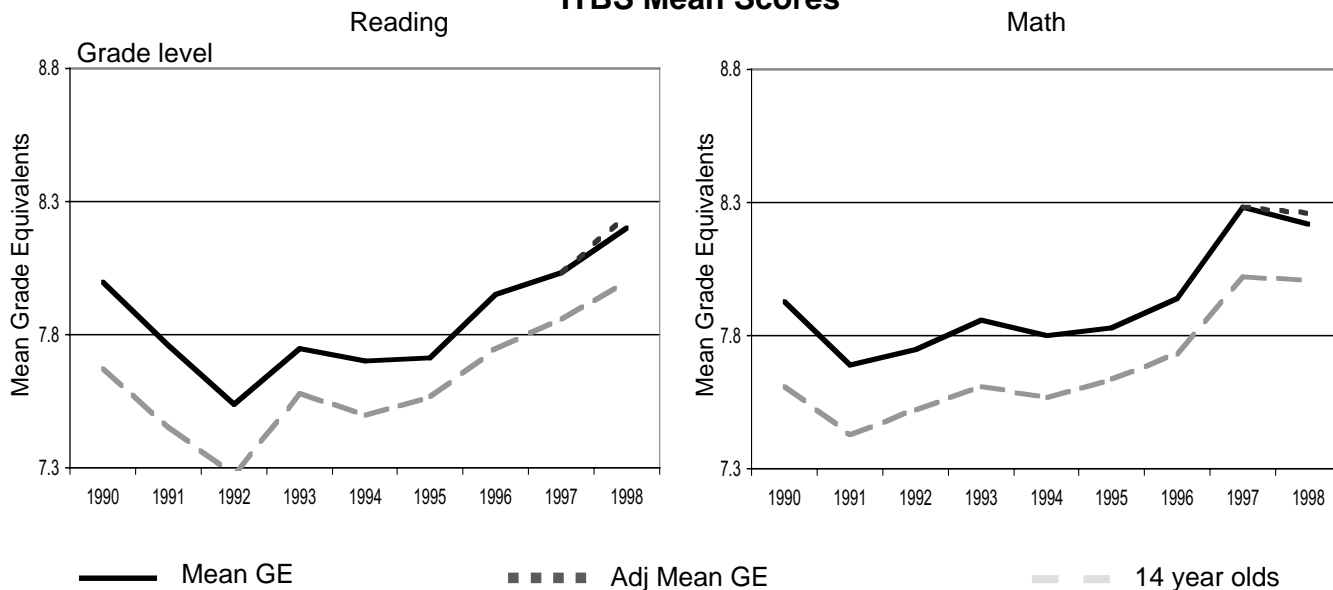


Grade Seven/13-Year-Old Students. Like fourth grade, seventh-grade scores are also negatively affected by the adjustments. Compared with the 1997 scores, adjusted 1998 scores are lower in reading but higher in math.

Figure 11

Figure 12

Eighth Grade/Fourteen Year Olds ITBS Mean Scores



Grade Eight/14-Year-Old Students. Adjusted reading and math scores in 1998 are higher than unadjusted scores. Adjusted 1998 reading scores are higher than 1997 scores; adjusted 1998 math scores are slightly lower than 1997 scores.

Figure 13

Mean ITBS Grade Equivalent Scores by Grade

Reading Comprehension

Grade	1990	1991	1992	1993	1994	1995	1996	1997	1998	1998 Adjusted
3	2.98	3.04	3.08	2.97	3.03	2.96	3.05	3.00	3.17	3.26
4	3.84	3.75	3.94	4.07	3.96	4.05	4.00	4.20	4.33	4.14
5	4.85	4.66	4.85	5.05	5.04	5.03	5.12	5.23	5.26	5.26
6	5.80	5.62	5.61	5.86	5.87	5.83	6.02	6.05	6.20	6.30
7	7.10	6.79	6.62	7.07	6.81	7.04	6.92	7.22	7.35	7.12
8	8.00	7.76	7.54	7.75	7.70	7.71	7.95	8.03	8.20	8.25

Math Total

Grade	1990	1991	1992	1993	1994	1995	1996	1997	1998	1998 Adjusted
3	3.27	3.32	3.25	3.42	3.38	3.39	3.45	3.35	3.61	3.65
4	4.18	4.15	4.10	4.27	4.20	4.23	4.24	4.45	4.60	4.45
5	5.08	5.03	5.12	5.15	5.15	5.14	5.22	5.33	5.40	5.40
6	6.09	6.04	6.03	6.25	6.13	6.23	6.24	6.47	6.47	6.54
7	7.03	6.91	6.98	6.99	6.86	6.96	6.94	7.16	7.39	7.22
8	7.93	7.69	7.75	7.86	7.80	7.83	7.94	8.28	8.22	8.26

Scores calculated by the Consortium on Chicago School Research.

1998 adjustments are made by removing retained students from 3rd, 6th, and 8th grades, and then including them in 4th and 7th grades (not 9th).

Figure 14

Mean ITBS Grade Equivalent Scores by Age Group

Reading Comprehension

Age Group	1990	1991	1992	1993	1994	1995	1996	1997	1998
Nine Year Olds	2.99	2.99	3.01	2.89	2.93	2.85	2.95	2.90	3.10
Ten Year Olds	3.82	3.77	3.89	3.92	3.82	3.89	3.86	4.01	3.97
Eleven Year Olds	4.76	4.61	4.80	4.95	4.89	4.86	4.93	5.04	5.05
Twelve Year Olds	5.71	5.55	5.59	5.87	5.81	5.74	5.86	5.91	6.08
Thirteen Year Olds	6.84	6.60	6.51	6.92	6.75	6.90	6.80	6.99	6.99
Fourteen Year Olds	7.67	7.45	7.27	7.58	7.50	7.57	7.75	7.86	8.00

Math

Age Group	1990	1991	1992	1993	1994	1995	1996	1997	1998
Nine Year Olds	3.32	3.31	3.20	3.34	3.30	3.30	3.36	3.27	3.50
Ten Year Olds	4.14	4.14	4.07	4.17	4.08	4.11	4.14	4.29	4.32
Eleven Year Olds	5.01	5.00	5.06	5.10	5.02	4.99	5.06	5.18	5.22
Twelve Year Olds	5.97	5.92	5.98	6.16	6.03	6.06	6.06	6.27	6.30
Thirteen Year Olds	6.87	6.79	6.86	6.93	6.83	6.90	6.85	7.04	7.11
Fourteen Year Olds	7.61	7.43	7.52	7.61	7.57	7.64	7.73	8.02	8.01

Scores calculated by the Consortium on Chicago School Research.

Figure 15

Percent of Students Scoring At or Above National Norms

Reading Comprehension

Grade	1990	1991	1992	1993	1994	1995	1996	1997	1998	1998 Adjusted
3	18.6	20.7	23.8	21.9	22.5	21.6	22.4	20.3	25.2	28.5
4	16.9	18.1	22.1	25.7	25.5	25.0	26.7	28.7	35.6	31.2
5	22.1	20.3	23.1	25.6	28.8	25.4	31.1	33.4	34.6	34.6
6	21.7	19.9	20.3	25.5	27.6	25.4	31.8	31.9	35.3	38.2
7	28.1	24.9	23.4	31.8	26.3	31.6	29.0	35.3	38.0	34.3
8	28.2	28.0	21.8	30.3	28.5	29.7	33.4	32.0	38.3	39.8
All Grades	22.3	21.8	22.4	26.8	26.5	26.5	29.1	30.3	34.3	34.4

Math Total

Grade	1990	1991	1992	1993	1994	1995	1996	1997	1998	1998 Adjusted
3	27.2	30.3	27.3	35.5	33.9	34.6	36.6	32.5	43.0	44.7
4	24.2	26.3	25.0	30.9	28.5	30.2	30.6	36.6	42.3	37.6
5	23.9	24.8	26.1	28.9	28.0	28.6	30.8	34.9	36.2	36.2
6	27.6	27.1	26.9	34.8	30.7	34.3	34.2	40.8	41.3	43.9
7	25.4	26.4	26.9	26.7	24.1	27.0	26.3	33.3	37.5	34.3
8	27.3	25.1	25.8	23.6	23.1	24.4	27.3	36.9	34.9	36.1
All Grades	25.9	26.6	26.3	30.1	28.0	29.8	31.0	35.9	39.3	38.7

Scores calculated by the Consortium on Chicago School Research.

1998 adjustments are made by removing retained students from 3rd, 6th, and 8th grades, and then including them in 4th and 7th grades (not 9th).

Endnotes

¹See Richard L. Allington and Anne McGill-Franzen, "Unintended Effects of Educational Reform in New York, Educational Policy 6:4 (December 1992): 394-414.

²The numbers of retained students whose scores were separated are: grade three, 4,120 in reading and 4,105 in math; grade six, 2,614 in reading and 2,596 in math; grade eight, 1,116 in reading and 1,106 in math. Additional students were also retained in these grades, but their scores are not included in public reporting and are therefore excluded from this analysis.

³Anthony S. Bryk, Yeow Meng Thum, John Q. Easton, Stuart Luppescu, *Academic Productivity of Chicago Public Elementary Schools* (Chicago: Consortium on Chicago School Research, 1998).

⁴See the technical appendix on our website, <http://www.consortium-chicago.org> for more information about the adjustment.

This data brief reflects the interpretations of the authors. Although the Consortium's Steering Committee provided technical advice and reviewed an earlier version of this brief, no formal endorsement by these individuals, their organizations, or the full Consortium should be assumed.