

COMMENT ON
Update: Ending Social Promotion

September 2000

Melissa Roderick, Jenny Nagoaka,
Jen Bacon, John Q. Easton
Consortium on Chicago School Research

**NEW DATA ABOUT CHICAGO'S GRADE
RETENTION PROGRAM PROVIDES FURTHER
PROOF THAT NEITHER RETENTION NOR
SOCIAL PROMOTION WORKS**

**Research-Based Alternatives to Both Retention
and Social Promotion Should Be Carried Out**

Prepared by
Donald R. Moore, Ed.D.*
Executive Director
Designs for Change

Additional Comment by
Professor Robert Hauser, Professor of Sociology, University of
Wisconsin; Chair, Committee on Appropriate Test Use,
National Research Council, National Academy of Sciences

September 7, 2000

Additional Information about Retention
Available at <www.designsforchange.org>

*The author is a member of the Steering Committee of the Consortium on Chicago School Research. The views presented here solely represent those of the author, and no endorsement by any other members of the Steering Committee or Consortium staff should be assumed.

TABLE OF CONTENTS

1. Section 1. Consortium <i>Update</i> Study Provides Critical New Evidence about Effectiveness of Chicago’s Closely-Watched Program to “End Social Promotion” ..p. 2	
2. <i>Update</i> Provides Additional Proof that Retention Does Not Benefit Retained Students.....	
.....p. 4	
3. Major Increase in Retention at Other Elementary Grades.....p. 6	
4. Further Evidence about the Dismal Achievement of Chicago’s Retained Students Should Lead to Drastic Changes in the Retention Component of the Program.....p.7	
5. Other Alleged Program Benefits Are Questionable.....p. 9	
Endnotes.....	
.....p. 14	
Tables.....	
.....p. 16	
Comment by Professor Robert M. Hauser.....p. 23	

SECTION 1. CONSORTIUM *UPDATE* STUDY PROVIDES CRITICAL NEW EVIDENCE ABOUT EFFECTIVENESS OF CHICAGO'S CLOSELY- WATCHED PROGRAM TO "END SOCIAL PROMOTION"

Background

In December 1999, researchers from the Consortium on Chicago School Research released the first independent study of Chicago's closely-watched effort to "end social promotion" (*Ending Social Promotion: Results from the First Two Years*).¹

The major focus of the initial Consortium study was to analyze the impacts of Chicago's test-based retention policy on students at grade three, six, and eight who faced retention in fall 1997.

Donald Moore of Designs for Change and Professors Gary Orfield of Harvard, Robert Hauser of the University of Wisconsin, and Ernest House of the University of Colorado prepared detailed critiques of the Consortium study (*Chicago's Grade Retention Program Fails to Help Retained Students: Better Alternatives Exist to Chicago's Costly Mistake*).² They highlighted important study findings, while taking issue with some of the Consortium study's interpretations of data and recommendations.

In particular, Moore, Orfield, Hauser, and House concluded that the retention portion of Chicago's program was harming thousands of children and should be radically restructured. The Consortium found that students who attended an intensive test prep summer school and then repeated their grade achieved no better than similar low-achieving Chicago students who were "socially promoted" in 1995 and were typically given no special help.³ (Social

promotion simply allows students to automatically move through school with their age group, whether or not they master the skills and content appropriate for their grade.)

Moore and the other commentators pointed out that the Chicago results were consistent with three decades of research which shows that retained students do not achieve significantly better in the long-term than similar low-achieving students who are simply socially-promoted and that retained students are much more likely to drop out later on.⁴ These conclusions, they pointed out, have held true even when retained students were provided with a different educational program in the year that they repeated a grade and substantial extra help.⁵

In contrast, the initial Consortium study concluded that "It is still quite early in this evaluation to make statements about whether the policy is working."⁶

Current *Update* Study

In September 2000, the Consortium released *Update: Ending Social Promotion*.⁷ This current *Update* study:

- Provides an **additional year** of information about the achievement of students who, in **spring of 1997**, faced Chicago's retention and promotion policy at the third, sixth, and eighth grade levels. The *Update* documents their progress through spring of 1999.
- Provides the first research analysis concerning the achievement of students who faced Chicago's retention policy at the third, sixth, and eighth grade levels in **spring of 1998**.
- Provides data about students who were retained in kindergarten, first, and second grade under the retention policy.

- Provides data about the dropout rates for eighth grade students retained under the retention policy, as compared with other eighth grade students.

How Chicago's Retention Plan Works

In the 1996-97 school year, Chicago initiated a massive effort to improve student achievement on standardized tests through the use of retention (requiring students to repeat a grade) and the threat of retention. In third, sixth, and eighth grade, students who failed to achieve a minimum test score in spring 1997 on the Iowa Test of Basic Skills in either reading or math or both were required to attend summer school (called the Summer Bridge program). At the end of Summer Bridge, they again took the Iowa Test in reading and math. If they still did not achieve the minimum test score in the subject or subjects they had failed in the spring, they were required to repeat their grade (that is, they were "retained" or "flunked") — unless they received a "waiver" from the central administration.

In addition to establishing specific test-based criteria for promotion at third, sixth, and eighth grade, Chicago's retention/promotion policy generally discouraged "social promotion" at all grades. (The Consortium *Update* and this Comment provide the first data on the extent of retention in kindergarten and at grades 1, 2, 4, 5, and 7.)

An Intensive Focus on Iowa Test Preparation Has Resulted from the Retention Policy

In the regular instructional program for all students, most Chicago elementary schools now devote a substantial portion of school time, beginning in March, to test preparation drills for the Iowa Test,⁸ which include such curriculum materials as *Test Best on the ITBS* and practicing on old versions of the Iowa Test.

For those students who fail to meet the cutoff score in the spring at grades three, six, and eight, the entire Summer Bridge program focuses on preparing students to pass the particular version of the Iowa Test for their grade level. The "Teacher Handbook" for the Summer Bridge program unabashedly describes the curriculum as a prep course to pass a particular level of the Iowa Test. The Introduction states:

An item analysis of those competencies essential to the achievement of success on the Reading/Language Arts and Mathematics ITBS and the Tests of Achievement and Proficiency was studied. The Summer Bridge Program's Teacher Handbook emphasizes the skills assessed in those subject areas.⁹

Even though testing experts conclude that such intensive test preparation invalidates the interpretation of test gains,¹⁰ every aspect of the curriculum is geared towards the Iowa Test. For example:

- The Summer Bridge handbook spells out 37 "process skills" assessed by the ITBS reading and math tests, as well as the percentage of questions at grades 3, 6, and 8 that are devoted to a particular process skill (for example, 8% of the reading items on the eighth grade ITBS require the student to "infer traits, feelings, or motivations of characters").

- The "Teacher Handbook" then presents a day-by-day hour-by-hour scripted curriculum of teacher-directed activities that provide virtually no room for teacher deviation. Teachers are only provided with two days of in-service to prepare to teach this curriculum. Proctors monitor teachers to make sure that they stay on the proper lesson. The content of the scripted curriculum is focused on the 37 process skills, according to the ITBS emphasis on them at a particular

grade level. Further, the process skills that are the focus of each lesson are listed at the beginning of each lesson.

Further, Designs for Change staff have been told by teachers and principals at a number of schools with a significant enrollment of retained students that the curriculum recommended by central office staff for the regular school year for retained students is to use the Summer Bridge curriculum as the core of the instructional program, since the Summer Bridge curriculum reflects “what the students will be tested on.”

Chicago’s Retention Program is Extremely Expensive

Based on an analysis of the Chicago school system’s budget, the cost of the Summer Bridge Program and extra teachers and tutors for retained students was \$86 million in the 1999-2000 school year.¹¹ In addition, it costs \$7,000 for each retained student for an extra year of school, so that retaining 10,000 students costs \$70 million per year. Thus, a reasonable estimate of the annual expense for the retention program is \$150 million per year.

Given this huge cost, it is critical that the effectiveness of the retention program is clearly documented and that alternative strategies for using these funds be considered.

Key Arguments on Which Chicago’s Retention Program is Based

One basic argument that Chicago school officials advanced in support of the policy was the common sense notion that low-achieving students will “catch up” and be prepared to succeed in the rest of their school experience, if they are given extra instruction in summer school and in the grade they repeat.¹²

A second argument made by Chicago school system leaders was that the “threat of retention” would motivate all students

to work harder, and that more students would meet the cutoff scores on the Iowa Test, either when they took the test in the spring or when they had completed a test prep summer school.

This Consortium’s *Update* and other data recently obtained by Designs for Change provide important new information that can be used to evaluate these claims.

SECTION 2. UPDATE PROVIDES ADDITIONAL PROOF THAT RETENTION DOES NOT BENEFIT MOST RETAINED STUDENTS

1997 Retained Students Are Still Achieving No Better than Socially Promoted Students

One key focus of the *Update* is to continue to track the achievement of students who were retained in fall 1997, as compared with similar low-achieving students who were “socially promoted” in 1995. The *Update* presents Iowa Test reading results for third and sixth graders over a three-year period, adding an additional year to the results presented in the original Consortium study.

The Consortium’s findings are initially presented in terms of “grade equivalents,” the standard that is used in Chicago to decide whether students are promoted or retained. (In Chicago, student scores on the Iowa Test are presented as “grade equivalents — for example, 6.8 means “sixth year, eighth month).

As shown in Table 1 of this Comment:

- Third graders who were socially promoted in 1995 were scoring two months higher in second grade than third graders who were retained in 1997. Yet three years later, socially promoted students were scoring seven months higher than retained students (see Table 1).¹³

- Sixth graders who were socially promoted from sixth grade in 1995 were scoring one month higher in fifth grade than sixth graders who were retained in 1997. Yet three years later, socially promoted students were scoring five months higher than the retained students (see Table 1).¹⁴

Thus, the researchers found that results based on “grade equivalent scores” indicated that socially promoted students were doing better than retained students at both grades 3 and 6. The researchers carried out further analyses using a different standard of measurement for comparing the test scores of retained and socially promoted students (called the Rasch logit model).¹⁵ The *Update* states that this model provides a more accurate measure of student gains than grade equivalent scores do.

After carrying out an analysis of third and sixth grade reading test scores using the Rasch logit method, the Consortium concluded that:

- Low-achieving third graders who were socially promoted in spring 1995 made the same three-year test score gains as third grade students who were retained in spring 1997.
- Low-achieving sixth graders who were socially promoted in spring 1995 still made greater three-year test score gains than sixth grade students who were retained in spring 1997.

Two points should be underscored in reviewing the comparison of the progress of retained and socially promoted students presented in the *Update*.

- First, the progress of the 1995 socially-promoted students equaled or exceeded the progress of the 1997 retained students, even though the socially-promoted students were typically passed to the next grade without attending summer school or receiving any other subsequent special programs. In contrast, the retained

students typically attended a test prep summer school (once and sometimes twice), repeated their grade in school, and received extra tutoring.

- Second, both socially promoted and retained students continued to **achieve very poorly**. Both retained third graders from 1997 and socially promoted third graders from 1995 were far below the national average for fifth grade two years later (see Table 1). The pattern for sixth graders is the same: both retained sixth graders from 1997 and socially promoted sixth graders from 1995 were far below the national average for eighth grade two years later (see Table 1).

Neither retention nor social promotion are effective. However, there are other effective research-based alternatives that can replace both retention and social promotion (as discussed on page 7).

Results for Retained Students Improved Little Between 1997 and 1998

When the initial Consortium report on the impact of retention showed poor results for retained students, school system leaders said that there would be a major improvement when 1998 results were available. However, as the *Update* shows, 1998 retained students were still achieving very poorly. The *Update* states:

Students who were retained in 1998 had an extra chance to meet the test cutoffs in January, and approximately one-quarter met the cutoff on this new testing date. Even with this extra chance, however, the performance of retained students was only slightly better than in 1997.¹⁶

The small improvement in the test performance of students who were retained in 1998, as compared with 1997, is reflected in Table 2, which is based on data presented in the *Update*.¹⁷

Table 2 indicates that even with the opportunity to attend summer, to repeat a grade with extra help, to take the Iowa Test an additional time at mid-year, and (for some retained students) to attend summer school a second time, only 53% of third graders, 55% of sixth graders, and 38% of eighth graders met the minimum cutoff score required for promotion.

Thus, even after extensive additional help at a cost of about \$150 million per year, 47% of retained third graders, 45% of retained sixth graders, and 62% of retained eighth graders failed to meet minimum test score cutoffs for promotion.

High Dropout Rate for Retained Eighth Graders

The *Update* documents a **29% dropout rate** for 1997 eighth graders who were retained in eighth grade or in a Transition Center — two years after these students were retained.¹⁸ This observed dropout rate is consistent with three decades of studies that indicate that retained students are more likely to drop out than similar low-achieving students who are not retained.¹⁹

In addition to the dropout rate for retained students, the number of students officially listed as “transfers” needs to be scrutinized carefully, as well as the status of retained eighth graders who are still in school (for example, the percentage of retained students who made a successful transition to a regular high school and are still enrolled). We have asked the Consortium for these data, and they have agreed to carry out the necessary analyses.

SECTION 3. MAJOR INCREASE IN RETENTION AT OTHER ELEMENTARY GRADES

The *Update* documents a substantial increase in the number of students

retained at kindergarten, first, and second grades.²⁰ The *Update* concludes that:

- Increased retention at these early grades is a significant contributor to the fact that there was a decline in the percentage of low-achieving third graders in 1998 and 1999, as compared with 1997. Observed gains in the percentage of third graders who reach the cutoff score on the Iowa Test in 1998 and 1999 need to be assessed in light of this finding.²¹
- Students who were retained in kindergarten, first grade, and second grade performed poorly in third grade. For example, “In 1999, 33% of third graders who had been retained the year before, in second grade, and were included under the policy were retained again.”²²

As the main author of the *Update* has recently stated elsewhere, “No research says that early grade retention is good for kids.”²³

Designs for Change has asked the Consortium for grade-by-grade systemwide data concerning the numbers of students retained at **all elementary grade levels** (grades K-8) from 1995 to 1999. Table 3 presents these data for students who were retained in a particular grade for the first time:

- At those grades where students did not have to meet a particular minimum test score on the Iowa Test to be promoted (kindergarten and grades 1, 2, 4, 5, and 7), the number of students who were retained typically doubled or tripled from 1995 to 1999, reflecting the school system’s general support for increased student retention.
- In fall 1999, 2,534 first graders and 1,498 graders were retained. More first graders were retained in 1999 than eighth graders (1,794).

Overall, the number of students retained at kindergarten and grades 1, 2, 4, 5, and 7 was about two-thirds as many as the number of third, sixth, and eighth graders retained. Table 4 shows that:

- The three-year total for 1997, 1998, and 1999 of kindergarten, first, second, fourth, five, and seventh graders retained was approximately 20,000.
- The three-year total for 1997, 1998, and 1999 of third, sixth, and eighth graders retained was approximately 34,000.

Thus, data about retentions at kindergarten and grades 1, 2, 4, 5, and 7 underscore the fact that the Chicago retention policy has a further negative impact that has previously received no public attention: large numbers of elementary students are being retained at other grades besides 3, 6, and 8.

The large number of students retained at kindergarten and grades 1, 2, 4, 5, and 7 is alarming given (1) the negative impact of retention on achievement at grades 3, 6, and 8 (even when students are receiving summer school and other extra help), (2) the Consortium's negative findings about the achievement of students who are retained in first and second grade, and (3) the increased dropout rate for retained eighth graders.

SECTION 4. FURTHER EVIDENCE ABOUT THE DISMAL ACHIEVEMENT OF CHICAGO'S RETAINED STUDENTS SHOULD LEAD TO DRASTIC CHANGES IN THE RETENTION COMPONENT OF THE PROGRAM

The *Update's* additional evidence about the negative impact of Chicago's retention program on retained students should result in its immediate restructuring. And

other school districts should no longer view Chicago as a model for improving the achievement of low-income students.

The Consortium's *Update* is limited to stating empirical findings and does not include any policy recommendations. However, the principal author of the Consortium study, Melissa Roderick, stated in a comment about Chicago's retention program recently published in the *American School Board Journal* that:

The effect of retention on these kids seems to be very decimating," Roderick says, explaining that children who are sent back to repeat a grade have a much higher dropout rate than children who progress with their classmates. "This is just a disaster, to be quite honest," she says. "And it's just the beginning of a disaster, because now we're seeing all of these first- and second-graders who are being retained." (Contrary to what many teachers and parents think, Roderick says, "no research says that early-grade retention is good for kids."²⁴

Designs for Change agrees with the *Update's* evidence about the severe negative impact of Chicago's retention program, which have been so eloquently summarized by the *Update's* principal author.

In addition, we agree with the recommendation that Professor Gary Orfield of Harvard made in his comments contained in the Designs for Change Rejoinder to the initial Consortium study of the impact of Chicago retention:

The problem is that there is massive evidence that this is a misbegotten policy, and tens of thousands of students' futures are at stake now. We should stop this policy now before there is irreparable damage to a generation of Chicago school children.²⁵

Research-Based Alternatives to Both Social Promotion and Retention Should Be Carried Out

Those who criticize the school system's retention/promotion policy have an obligation to spell out workable research-based alternatives.

Most people assume that the only two policy options that a school system can carry out are (1) social promotion (defined as allowing students to move through school automatically with their age group whether or not they master skills and academic content appropriate to their grade) or (2) retention.

Both retention and social promotion are **failed policies**. Fortunately, there are a number of alternative approaches, strongly supported by research, that will enable urban school systems like Chicago both to assist a high proportion of their students to graduate from secondary school and to master academic skills needed for future success. These alternatives represent much more productive strategies for spending the massive amounts now being wasted on retention.

I have briefly summarized the characteristics of these workable alternatives below.

Investing in Prevention. Several highly effective alternatives to both social promotion and retention are preventative; their aim is to help students succeed consistently in the regular school program, so that educators and families do not have to face the difficult decision about whether to promote or retain a low-achieving student.

These include, for example:

- High quality early childhood education.
- Fundamental school restructuring by changing teaching and learning from 9

a.m. to 3 p.m. during the regular school year.

- Preparing all teachers to become effective reading teachers.

As a Last Resort, Promotion with Special Help Is Preferable to Retention.

In the conclusions presented in a previous study of dropouts in Fall River, Massachusetts that Melissa Roderick, the *Update*'s principal author, published in 1993, she stated the following:

Research on grade retention has found that **promotion with remediation** provides more short-term academic benefits to youths than either retention alone, retention with remediation, or promotion alone (Karweit, 1991; Smith and Shepard, 1989). The results of the Fall River study lend additional support to the conclusion that promotion with remediation is the appropriate alternative (emphasis added).²⁶

Designs for Change agrees with this conclusion and advocates a fundamental departure from the current retention process. It is more effective to promote students but give them intensive special help, rather than to retain them. The most careful study of this option indicates that low-achieving students who were promoted and given special help achieved better than either retained students who were provided with special help or students who were promoted without special help.²⁷ Further, promoted students who receive special help avoid the dangers of increased dropout that have been repeatedly documented in previous research about retention.

Further, providing promotion with special help to all low-achieving students addresses one of the problems with the current Chicago initiative documented in the *Update*: Low-achieving students who obtain waivers to be promoted do not currently receive special help in the next school year. Low-achieving sixth graders

who received waivers did very poorly when they reached the eighth grade promotion gate.²⁸

A Comprehensive Alternative to the Current Retention Policy

Thus, Designs for Change advocates a comprehensive alternative to the current retention policy that includes the following key elements:

- Family choice about retention.
- Individual learning plans for all low-achieving students.
- Additional sustained help for low-achieving students, primarily in the regular classroom.
- Ongoing teacher education.
- Parent participation in shaping a child's learning program.

The appropriateness of family choice about retention deserves particular emphasis. The *Update's* additional negative findings about the impact of retention reinforce scores of studies indicating that retention is ineffective and dangerous. Just as physicians need parent permission to undertake a risky operation, educators should obtain parent permission to require a student to repeat a grade. After reviewing all the options with school staff, families should have the right to choose retention for their child coupled with special help, if they wish. However, the working rule should be that a low-achieving student will be promoted and given special help, unless their family specifically asks for the student to be retained.

Reinvest the Costs of Retention in Alternatives That Work

These alternatives will obviously cost a significant amount of money. However, the large investment that Chicago is now

making in retention (approximately \$150 million a year) can be invested in alternative strategies that have been shown through research to improve student achievement and graduation rates.

For example, an extra year of schooling for a retained students costs more than \$7,000 a year, so the cost of an extra year of schooling for 10,000 retained students just at grades three, six, and eight is \$70 million. These funds would pay the salaries and benefits of at least 1,000 additional teachers to provide extra help to low-achieving students.

SECTION 5. OTHER ALLEGED PROGRAM BENEFITS ARE QUESTIONABLE

Some elements of Chicago's current retention/promotion policy are clearly benefiting significant numbers of students. For example, providing more instructional time for low-achieving students during the summer is a sound approach supported by research, although the summer program would be more effective if it were based on individual assessment of student learning problems and individualized instruction, rather than a "one-size-fits-all" scripted curriculum.

However, both the Consortium's original study and its *Update* provide evidence suggesting that the Chicago retention initiative might have benefited two groups of students, because the "threat of retention" has motivated them to work harder and learn more:

- Students who were motivated to work harder during the regular school year by the threat of retention and thus passed the minimum cut-off score on the Iowa Test in May.
- Students who made test score gains during Summer Bridge and were promoted.

However, these claims need to be scrutinized carefully, since a great deal of evidence **casts doubt on them.**

What Caused Observed Systemwide Increases in Students Meeting Minimum Iowa Test Cut Scores?

Table 5 summarizes the percentage of students who met the cut score on the Iowa Reading Test at grades 3, 6, and 8 in 1995, 1996, 1997, and 1998, based on data presented in the *Update*.²⁹ For 1997, 1998, and 1999, the bars in the chart represent both the number of students who passed the cutoff score in reading at the end of the regular school year and the additional number who passed the cutoff score at the end of summer school.

Similar bar charts have been used by proponents of the retention policy to make such statements as “80% of students are benefiting from the retention policy.” However, closer examination of Table 5 shows that this statement is untrue, even if one accepts the gains at face value. In 1995, the Iowa Test had no consequences for schools or students; the results were not even reported in the newspapers. Instead, the state’s IGAP Test was the focus for school accountability.

Even in 1995, however, 55% of third graders, 63% of sixth graders, and 66% of eighth graders met the test cutoff minimums. Thus, even if one regards the results presented in Table 5 as valid indicators of progress, one must weigh this progress in light of **the increase** in the percentage of students who met the cutoff scores in 1995 versus 1999.

Further, by 1997, the Iowa Test had not only become the criterion for student promotion and retention, but also the dominant criterion for placing schools on probation, and a major criterion for principal evaluation. And as noted earlier, Chicago’s retention/promotion initiative has multiple components, and

no systematic evidence has been presented that “the fear of retention” is the cause of the year-to-year increases shown in Table 5.

In their earlier report, the Consortium authors laid out five possible reasons for improvements from 1995 to 1998 (including the possibility that students were simply being prepped for one particular test). The authors then stated “At this point, we do not have sufficient evidence to evaluate these competing hypotheses.”³⁰

In its Rejoinder to the Consortium’s initial study of Chicago’s retention program, Designs for Change has laid out a number of reasons that might account for the test score gains shown in Table 5. The Rejoinder also spelled out evidence that could help determine the extent to which these increased scores reflect genuine learning gains, as opposed to intense test preparation. Below, we highlight two major problems with the Chicago test prep and testing program that need to be taken into account before anyone can conclude that increased Iowa Test scores reflect genuine learning.

Teaching to the Test

Professor Robert Hauser of the University of Wisconsin chaired the Committee on Appropriate Test Use of the Board on Testing and Assessment of the National Research Council. The committee closely examined the Chicago testing and retention program, which is described in the Committee’s report *High Stakes: Testing for Tracking, Promotion, and Graduation*.³¹ Professor Hauser has identified basic flaws in claims about Chicago test score gains during the regular school year and during summer school:

The NRC Committee concluded that Chicago’s regular year and summer school curricula were so closely geared to the ITBS that it was impossible to distinguish between real subject mastery and mastery of skills and

knowledge useful for passing this particular test... The only apparent change is a summer-school effect, which combines regression to the mean with intensive test preparation — but which soon tapered off. Chicago’s scores also reflect the typical upward “test-score creep” associated with repeated use of the same examination, but that is no indication of successful reform. Use of an independent, external standard of academic achievement — not the ITBS — is essential to a valid evaluation — but the Chicago study includes no independent standard of achievement.³²

Professor Ernest House, who audited the evaluation of a massive retention program in New York City in the 1980s that mirrors Chicago’s program in many respects, echoes Professor Hauser’s statement when he says, “Since teaching to the test is a major strategy of the Chicago program, the critical question is whether students can demonstrate gains on tests that they are not being taught. So far, the evaluation has not addressed this issue.”³³

House and other researchers have pointed out numerous examples of schools and school systems that have made substantial year-to-year progress on a particular high stakes test that was given year after year (like the Iowa Test), but did not score comparably on another test covering similar topics.³⁴

Repeated Use of the Same Identical Versions of the Iowa Test

Another practice that is contributing to increased test scores from 1997 through 1999 is Chicago’s repeated use of identical forms of the Iowa Test. Designs for Change has recently gathered information about which specific forms of the Iowa Test have been used in Chicago during the last decade for midyear, spring, and summer testing. As shown in Table 6:

- Form K of the Iowa Test has been used six times in the last eight years, and four times in the last four years. Also, Form K was used after Summer Bridge in three successive summers: 1997, 1998, and 1999.
- Form L of the Iowa Test has been used seven times in the last seven years, and five times in the last three years.

Such repeated use of the same test gives teachers who administer the tests multiple opportunities to make copies, to jot down questions, or simply to remember questions. For example, each version of the three forms of the eighth grade Iowa Reading Test that has been used repeatedly in the last three years has 43 vocabulary words on it. If a teacher writes down the 129 vocabulary words on these three versions of the test and subsequently teaches them to his class, this practice completely undermines the validity of the test. Thus, the test **not secure**, and the results **cannot be trusted**.

Research elsewhere has documented that test scores improve from year to year when the same test is given repeatedly, but then drop when a new secure test is introduced.³⁵

The Greatest Good for the Greatest Number?

As noted above, many in the public have interpreted the Consortium’s findings as indicating that a relatively small number of students are potentially being harmed by the retention/promotion policy (i.e., those students who are retained), but that a much larger number of students are potentially benefiting because of the threat of retention.

However, in assessing whether the retention/promotion policy is achieving “the greatest good for the greatest number,” it is critical to pin down the number of students who are potentially benefiting or being harmed. One

approach to illuminating this issue is reflected in Table 7. In 1999:

- About 12,000 students in grades 3, 6, and 8 were retained in fall 1999.
- The number of students who met the minimum cutoff scores in the spring at grades 3, 6, and 8 increased by about 14,000 students between 1995 and 1999.
- The number of students who attended Summer Bridge and then met the test score cutoff was about 5,000.

These results **undercut the view** that the number of additional students who met the cutoff scores and were promoted **vastly outweighs** the number of students who were retained at grades 3, 6, and 8 in 1999.

Further, an analysis that tallies up winners and losers should take into account:

- Issues raised above about whether the test score gains of students who were promoted are genuine and can be attributed to the retention/promotion policy.
- The approximately 6,500 students who were retained in kindergarten and grades 1, 2, 4, 5, and 7 in 1999 (see Table 4).

The Threat of Retention as an Educational Policy: Some Ethical and Practical Dilemmas

The school system leadership has repeatedly stated that a central element of their strategy is to use the “threat of retention” to motivate students, and they point to increased gains during the regular year and during Summer Bridge reflected in Table 5 as evidence of the efficacy of this threat. However, as noted above, the authors of the initial Consortium report were uncertain that the observed gains

were genuine, or what may have caused them.³⁶

Table 7 reflects the fact that the increase in the number of students who met minimum test score cutoffs in 1999 versus 1995 at grades 3, 6, and 8 is only moderately larger than the number of students who were retained at these grades.

If the Consortium study had shown that retained students benefited from retention, “the threat of retention” might be considered a legitimate public policy tool for improving student achievement. Yet even if some students who pass the test cutoff scores are benefiting from the threat of retention, the devastating impact of retention on the students who are retained raises severe ethical and practical concerns about the use of retention as a threat.

It is **unethical** to use 12,000 students as “sacrificial lambs” to motivate 19,000 additional students to meet test score cutoffs. Consideration of a similar approach that might be applied in the medical treatment of children quickly indicates that this approach is unconscionable.

Further, even if one could ignore the unethical nature of harming some students to motivate others, the practical consequences of such a strategy should deter an urban community from pursuing it. If a total of 15,000 to 20,000 students are retained each year in grades kindergarten through 8 (with resulting low achievement and increased propensity to drop out), they will not disappear. Most are likely to become low-skilled disaffected adults, with little chance to get a job with a future or to pursue further education in the nation’s rapidly changing economy. The decision to retain them will cost the society far more money in the long run than it would cost to carry out the research-based alternatives to both retention and social promotion summarized earlier.

Since retention and social promotion are both harmful, it is time to pursue research-based alternatives to both these failed policies.

ENDNOTES

¹ Melissa Roderick, Anthony S. Bryk, Brian A. Jacob, John Q. Easton, Elaine Allensworth (1999, December) *Ending social promotion: Results from the first two years*. Chicago: Consortium on Chicago School Research. (Hereafter, "ESP.")

² Donald R. Moore (April, 2000, Updated). *Chicago's grade retention program fails to help retained students: Better alternatives exist to Chicago's costly mistake*. Chicago: Designs for Change..

³ *ESP*, p. 35.

⁴ Lorrie A. Shepard and Mary Lee Smith, eds. (1989). *Flunking grades*. London: Falmer Press; Jay Heubert and Robert Hauser, eds. (1999). *High stakes: Testing for tracking, promotion, and graduation*. Washington: National Academy Press.

⁵ *Ibid.*; Ernest R. House (1998, November). *The predictable failure of Chicago's student retention program*. Paper presented at the Conference on Rethinking Retention to Help All Students Succeed. Chicago. (available at www.dfcl.org).

⁶ *ESP*, p. 59.

⁷ Melissa Roderick, Jenny Nagoaka, Jen Bacon, John Q. Easton (2000, September). *Update: Ending social promotion: Passing, retention, and achievement trends among promoted and retained students 1995-1999*. Chicago: Consortium on Chicago School Research. (Hereafter, *Update*)

⁸ BetsAnn Smith (1998, December). It's about time: Opportunities to learn in the Chicago Public Schools. Chicago:

Consortium on Chicago School Research., p. 23.

⁹ Chicago Public Schools (1999). *Bridging the gap to the 21st century: Teacher handbook: Language Arts*. Chicago: Author.

¹⁰ Heubert and Hauser

¹¹ Duffrin, Elizabeth (2000, April). Numbers confirm good, bad of promotion policy. *Catalyst*, 11 (7): 1, 4-6..

¹² Grant Pick (1998, April). Student retention: Trying to succeed where others failed. *Catalyst* (Vol. XI, No. 7), pp. 1, 4-9.

¹³ The data in Table 1 are taken from Figure 13 of the *Update* (p. 17).

¹⁴ The data in Table 2 are taken from Figure 11 of the *Update* (p. 15).

¹⁵ *Update*, pp. 16, 18, and 19.

¹⁶ *Update*, p. 10.

¹⁷ *Update*, p. 11.

¹⁸ *Update*, p. 12.

¹⁹ Shepard and Smith; Heubert and Hauser; Melissa Roderick (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31(4), 729-759..

²⁰ *Update*, p. 7.

²¹ *Ibid.*

²² *Update*, p. 9.

²³ Lawrence Hardy (2000, September), The trouble with standards. *National School Board Journal* 187(9), p. 31.

²⁴ *Ibid.*

²⁵ Moore (2000, April), p. 27.

²⁶ Melissa Roderick (1993). *The path to dropping out: Evidence e for intervention*. Westport, CN: Auburn House., p. 159.

²⁷ Gaea Leinhardt (1980). Transition rooms: Promoting maturation or reducing education. *Journal of Educational Psychology*, 72(1): 55-61.

²⁸ *Update*, p. 21.

²⁹ *Update*, Figures 4, 5, and 8 on pp. 6, 8.

³⁰ *ESP*, p. 26.

³¹ Heubert and Hauser, p. 132.

³² Moore (2000, April), p. 29.

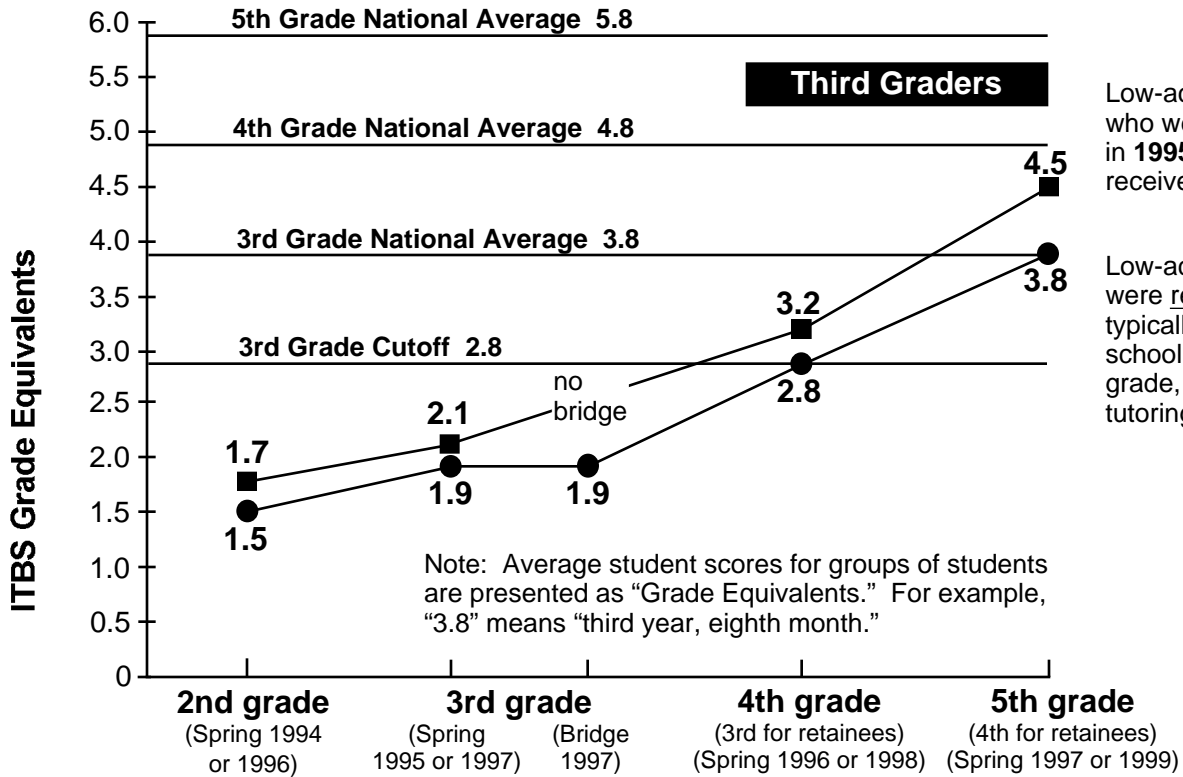
³³ *Ibid.*, p. 31.

³⁴ House; Daniel M. Koretz et al. (1991). The effects of high stakes testing on achievement: Preliminary findings about generalizations across tests. Paper presented at the Annual Meetings of the American Educational Research Association (Chicago, April 3-7).

³⁵ David Hoff (2000, January 26). Testing's ups and downs are predictable: Research shows cyclical pattern. *Education Week*, 24(2): 1; 12-13.

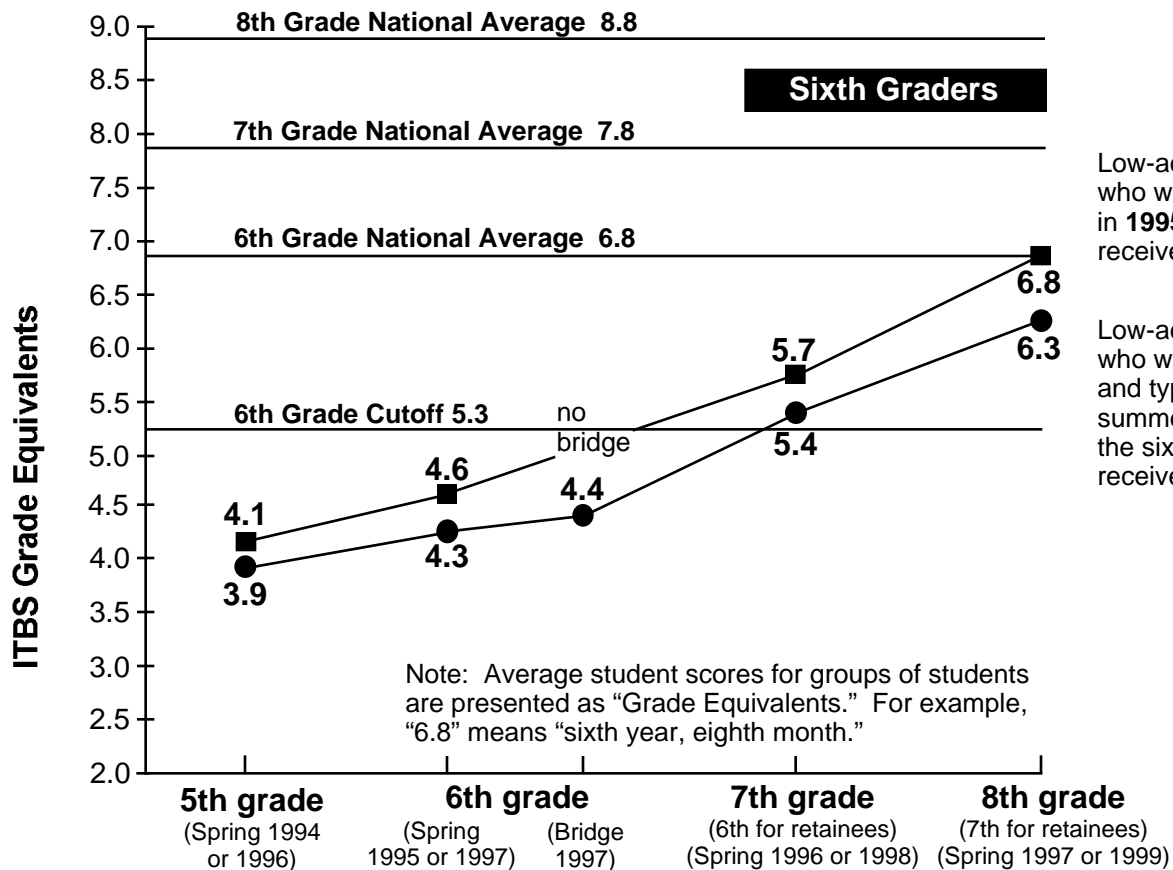
³⁶ *ESP*, pp. 26, 53).

TABLE 1. Three-Year Growth in Iowa Reading Test Scores for Retained and Socially Promoted Students of Third and Sixth Grades



Low-achieving students who were socially promoted in 1995 and typically received no extra help.

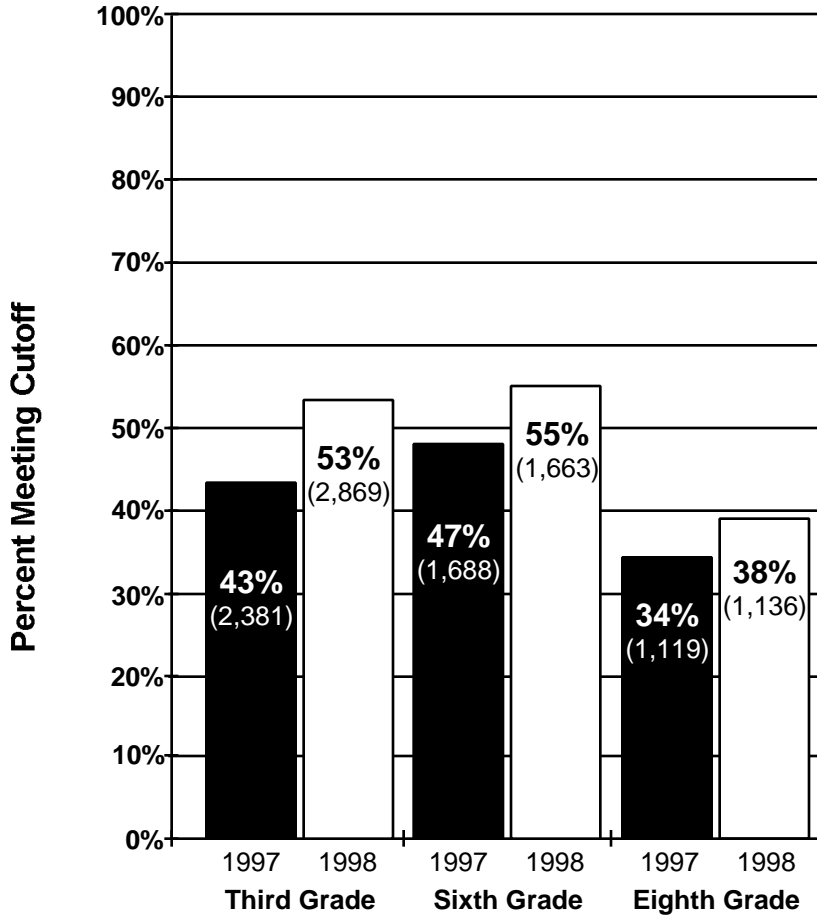
Low-achieving students who were retained in 1997 and typically attended summer school, repeated the third grade, and received extra tutoring.



Low-achieving students who were socially promoted in 1995 and typically received no extra help.

Low-achieving students who were retained in 1997 and typically attended summer school, repeated the sixth grade, and received extra tutoring.

TABLE 2. Percent of Students Retained in Fall 1997 and 1998 Who Met the Minimum Test Score Cutoff by August of the Year Following Their Retention



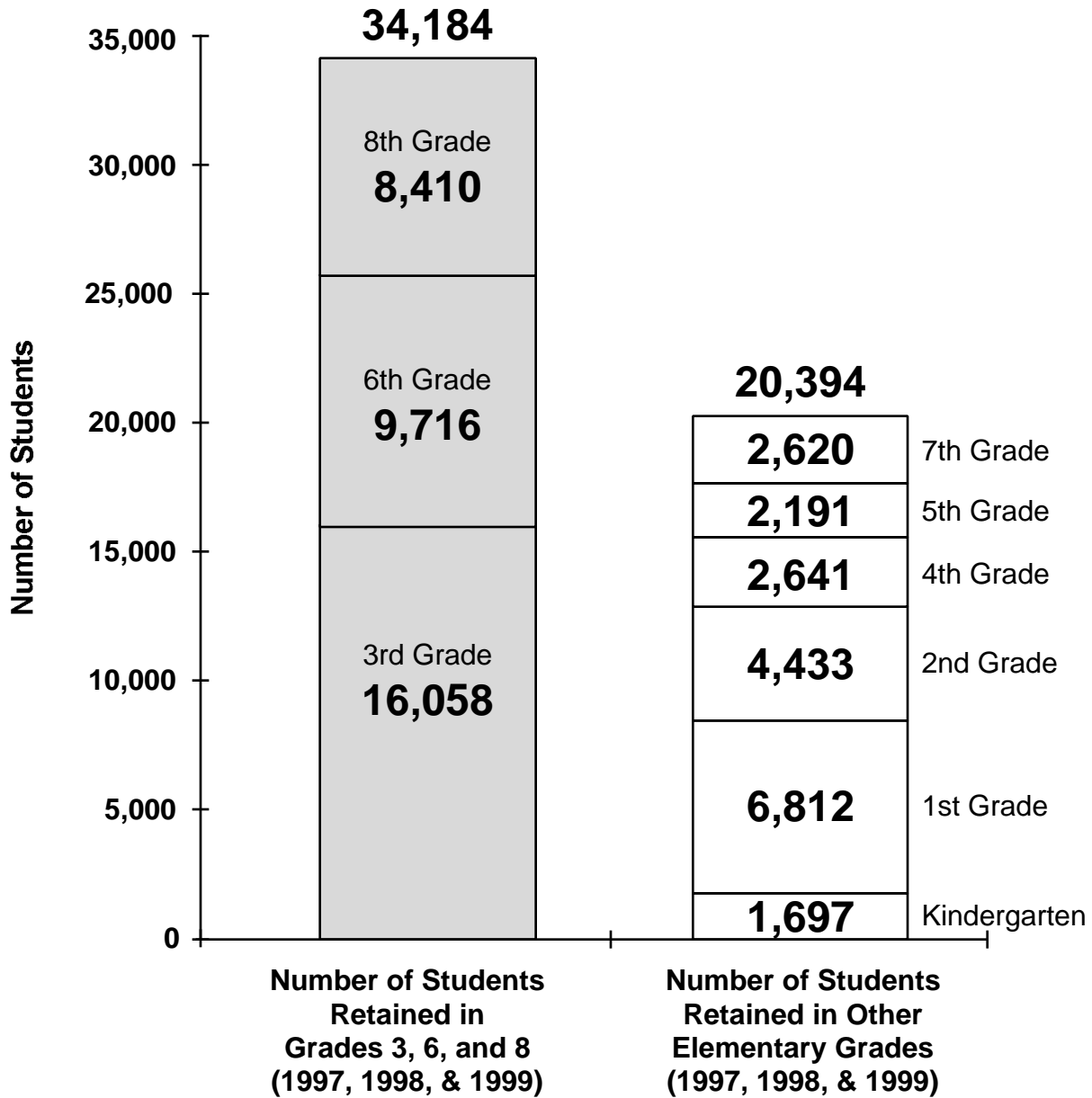
Note: These percentages reflect the total number of retained students who met the minimum test score cutoff at mid-year of their retention year (January), in spring of their retention year (April), or after attending Summer Bridge a second time (August).

**TABLE 3. Number of First Time Student Retentions in Grades K–8
between 1995 and 1999**

Grade	1995	1996	1997	1998	1999	Total
K	248	402	404	512	781	2,347
1	1,235	1,532	2,052	2,292	2,468	9,579
2	720	1,000	1,507	1,479	1,447	6,153
3	503	948	5,313	5,552	5,193	17,509
4	362	742	1,038	775	828	3,745
5	328	579	912	796	483	3,098
6	346	685	3,459	3,127	3,130	10,747
7	357	651	1,116	744	760	3,628
8	775	1,950	2,865	3,144	2,401	11,135
Total	4,874	8,489	18,666	18,421	17,491	67,941

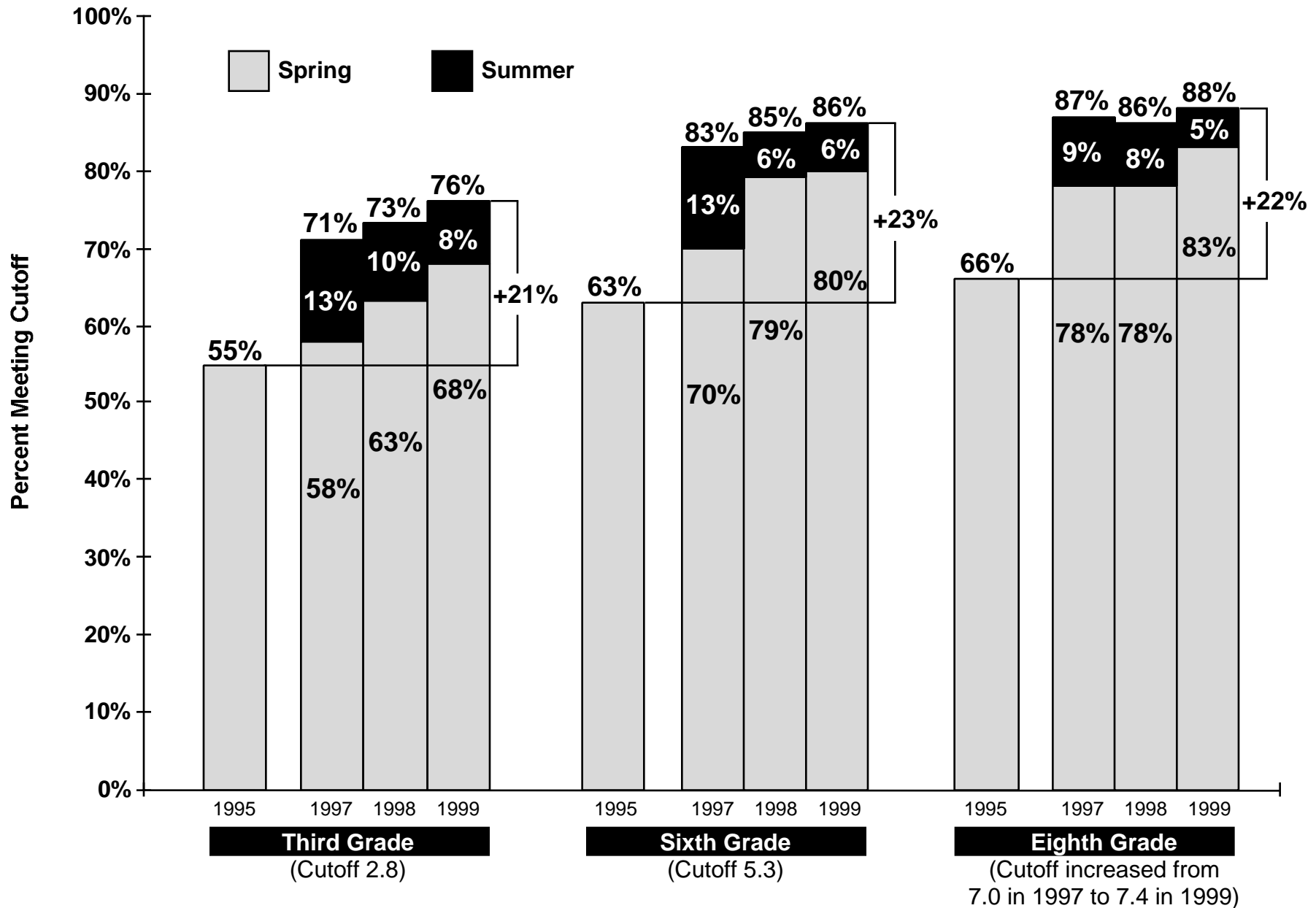
Note: Due to small differences in the methods for identifying numbers of retained students in data provided to Designs for Change (as contrasted with other Consortium studies), the figures shown in this table may vary slightly from results presented in other Consortium studies. Thus, figures in Table 3 should be considered as sound estimates of numbers of retained students.

**TABLE 4. Number of Students Retained in Grades K–8
Between 1997 and 1999**



Note: Totals shown in this table were obtained from Table 3.

TABLE 5. Percent of All Third, Sixth, and Eighth Graders Reaching the Iowa Test Cutoff in Spring and by the End of Summer Bridge: Years 1995, 1997-1999



Note: 1998 data provided separately by the Consortium on Chicago School Research.

Source: Consortium on Chicago School Research, *Update: Ending Social Promotion* (September 2000)

Prepared by Designs for Change
September 7, 2000

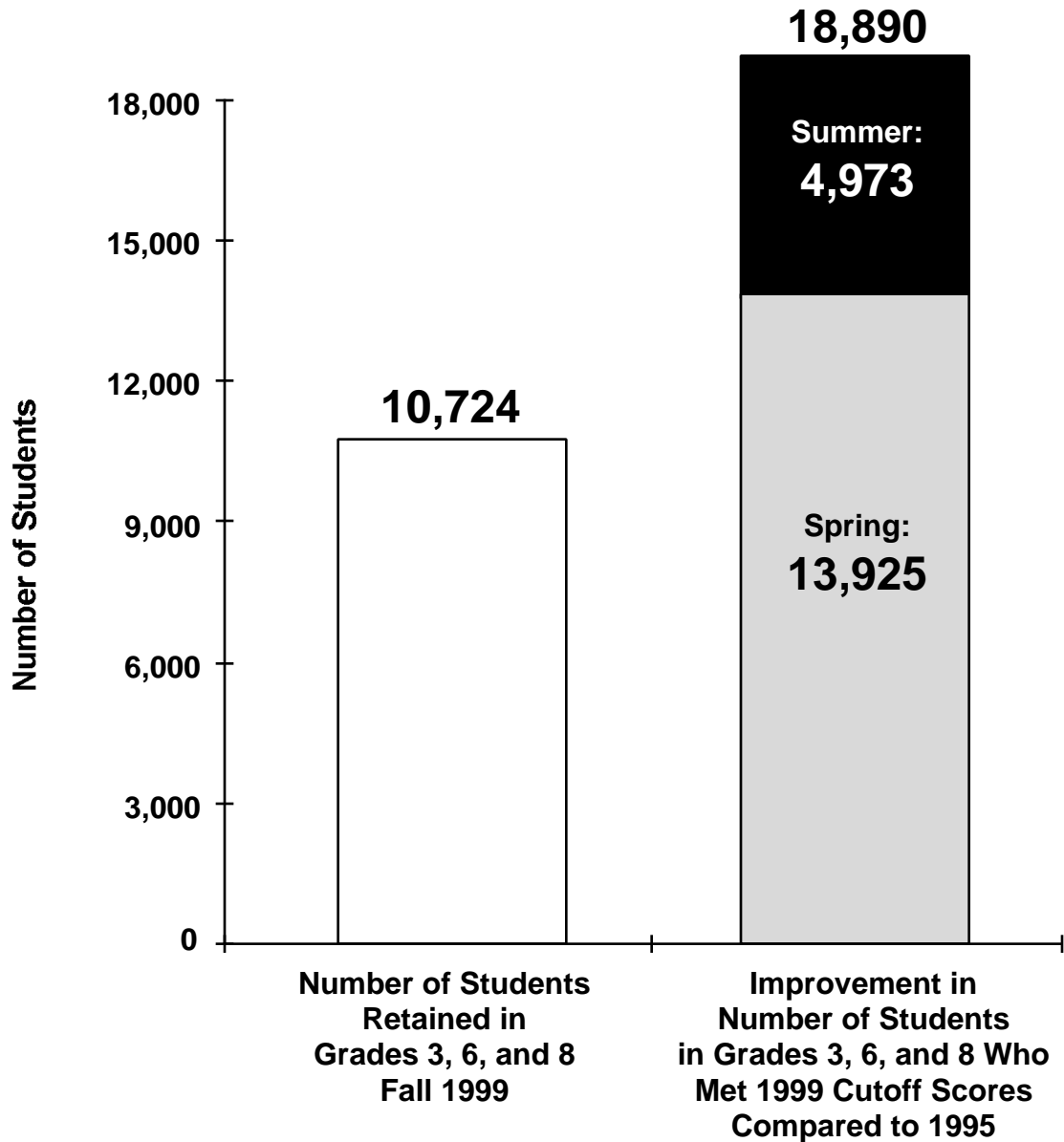
**Table 6. Form of the Iowa Test Administered
in Grades Three, Six, and Eight: Years 1990-2000**

Calendar Year	Mid-Year (January)	Spring (April)	Summer (August)
1990	—	CPS90 (J)	—
1991	—	CPS91 (G)	—
1992	—	CPS92 (H)	—
1993	—	CPS93 (K)	—
1994	—	CPS94 (L)	—
1995	—	CPS93 (K)	—
1996	—	CPS94 (L)	—
1997	—	CPS97 (M)	CPS93 (K)
1998	CPS94 (L)	CPS94 (L)	CPS93 (K)
1999	CPS94 (L)	CPS97 (M)	CPS93 (K)
2000	CPS94(L)	CPS93 (K)	CPS 94 (L)

Sources: Consortium on Chicago School Research and The Chicago Public Schools, Office of Accountability, Division of Student Assessment.

**Prepared by Designs for Change
September 7, 2000**

TABLE 7. Groups Potentially Impacted by the Promotion Policy in 1999



Sources: Consortium on Chicago School Research publications, *Ending Social Promotion* (December 1999), *Annual CPS Test Trend Review, 1999* (May 2000), Update: *Ending Social Promotion* (September 2000).

Analysis by Designs for Change
September 7, 2000

CENTER FOR DEMOGRAPHY AND ECOLOGY

The University of Wisconsin-Madison
1180 Observatory Drive
Madison, Wisconsin 53706

Robert M. Hauser
Vilas Research Professor of Sociology
(608) 262-2182, 262-8400 (FAX)
HAUSER@SSC.WISC.EDU

September 7, 2000

The Failure of Test-Based Retention in Chicago

I have read the December 1999 and September 2000 reports of the Chicago study carried out by the Consortium on Chicago School Research (CCSR), in light of the debate over ending “social promotion.” The latest, sad findings reconfirm and strengthen those of the first report.

- The combination of test-based retention with a narrowly test-based curriculum has no remedial educational value.
- Over a three-year period, students who have been retained gain no more on the ITBS than students who have passed promotional gates or who have failed summer examinations and been waived into the next grade. But those retained students lose a year of their lives, their retention adds to the cost of schooling, and they have a higher risk of dropping out before completing high school.
- Low-scoring eighth graders are more likely to drop out within two years under the new policy than under the policies of 1995, and dropout is especially high among low-scoring eighth graders who were retained or sent to transition centers in 1997.
- There are higher rates of passing the promotional gates at grades 3, 6, and 8, but they are no reason to celebrate. Gains have been very small between 1998 and 1999. Part of the gain in grade 3 is explained by more retention in grades 1 and 2. Part of the gain in all grades is due to increased exemption from the testing program. Moreover, fewer waivers have been issued after 1997, so retention rates remain very high.
- Most important, repeated use of the same forms of the same tests invariably leads to test-score inflation, and that may well account for the observed increase in pass rates.

The meager impact of the Chicago reforms is especially evident in comparisons between third-graders who failed the ITBS in the spring of 1997 and third-graders in the spring of 1995 whose scores were so low that they would have failed. (That is, I ignore the differences among students who passed after the 1997 summer bridge and who were retained or waived thereafter.) From Figure 2-9 of the first Consortium study, I calculated that the small initial difference between the two low-scoring groups (0.1 ITBS Grade Equivalents in the December 1999 report and 0.0 among continuing students in the September 2000 report) was unchanged a year later – after the 1997 students had gone through the first round of testing, the Summer Bridge program, and the following academic year. This finding is reinforced by the additional year of data: As shown by Figure 14 of the September 2000 report, the average three-year gain in GE is 2.7 among students who would have failed in 1995, and it is 2.7 among all students who failed the ITBS in the spring of 1997, regardless of subsequent participation in the summer bridge or later retention. That is, the overall test-score gain of initially low-scoring 3rd graders was essentially unaffected by selection after the 1997 summer bridge into groups that were passed, waived, or retained. The program had no

overall effect on test-score gains of low-scoring third-graders after three years, but we can expect that it will eventually increase school dropout among the retained students.

Much the same finding holds among students at the sixth grade level. There is virtually no difference after three years between the test-score gains of low-scoring sixth graders in 1995 and the aggregate of those who initially failed the ITBS in the spring of 1997 and were subsequently successful or unsuccessful in the summer bridge program. Again, the program had no effect on test-score gains of low-scoring sixth graders after three years, but we may expect it to increase school dropout.

Both among all third-graders and among all sixth-graders, there was a small overall increase in test-score gains between those who first completed the grade in 1995 and those who first completed the grade in 1997. But the gain is due entirely to the larger share of students who initially passed the ITBS. Yet we know that the increased passing rates are of doubtful validity because of the repeated use of the same forms of the same test after 1995.

In its 1999 report on high stakes testing, the National Research Council (NRC) stated a few basic principles of appropriate test use: (1) Any particular test has validity only in relation to specific uses; (2) tests are not perfect, but neither are the alternatives to tests; (3) no high-stakes educational decision about a test-taker should be made solely or automatically on the basis of a single test score; other relevant information should also be taken into account; (4) tests should be used for high-stakes decisions only after students have been taught the knowledge and skills on which they will be tested; (5) neither test scores nor any other kind of information can justify educational decisions that are not beneficial for students.

At the outset, the Chicago plan violated these basic principles. The Iowa Test of Basic Skills (ITBS) was not designed to be used as the sole criterion of grade-promotion decisions. Test reliability was not considered in setting arbitrary cut scores, and there was no second chance to take the spring exam. No other clearly specified criteria were used at first – though we now have learned from the Consortium study that thousands of students have actually been promoted, despite failing scores. The ITBS has no verified relationship to previously established curricular goals and standards of the Chicago Public Schools. In fact, the NRC Committee concluded that Chicago's regular year and summer school curricula were so closely geared to the ITBS that it was impossible to distinguish real subject mastery from mastery of skills and knowledge useful for passing this particular test. Again, this problem is exacerbated by Chicago's repeated use of exactly the same forms of the test.

This latest evaluation of the Chicago plan reconfirms that it has also failed the last of the NRC's criteria; there is no credible evidence of lasting educational benefit, either to those who passed or those who failed the ITBS. Use of an independent, external standard of academic achievement – not the ITBS – is essential to a valid evaluation – but the Chicago study includes no independent standard of achievement.

The failure of the Chicago plan will come as no surprise to anyone who has followed research on the consequences of grade retention. While retention sometimes leads to short-term

gains for retained students – and even these are lacking in Chicago – study after study shows that the costs of retention are high, both to school systems and to students, and that retention leads to lower academic achievement and greater high school dropout. And a worrisome level of school dropout is already signaled by the Chicago data. These findings reconfirm the failure of retention as shown in New York’s Promotional Gates program of the 1980s and by more recent studies in Baltimore and in Texas.

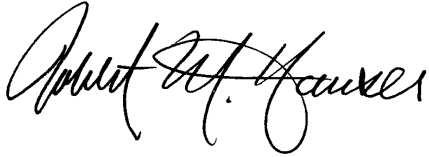
The Consortium reports merely tell us, again, what has happened in the past, not what need happen in the future:

- Chicago’s school leaders should ask themselves, what are the likely consequences, both immediate and in the long-term, of their well-intentioned efforts to raise educational standards?
- Chicago’s leaders need strong evidence that reforms will work before they put them in place on a large scale.
- Chicago’s leaders should maintain a commitment to measure reforms and their consequences as they take place. The Consortium study is a commendable move in this direction, but its value is reduced by the absence of an independent criterion of achievement and by other problems of research design and analysis.
- Chicago’s leaders should remember what the problem is as they seek to solve it. The problem is not social promotion; it is low academic achievement.

During 1998, I served as Chair of the Committee on Appropriate Test Use of the Board on Testing and Assessment at the National Research Council. The National Research Council is the operating arm of the National Academy of Sciences, which was chartered by Congress in 1863 to advise the government on matters of science and technology. I was elected to the Academy in 1984. The Committee on Appropriate Test Use prepared its report, *High Stakes: Testing for Tracking, Promotion, and Graduation* (National Research Council, eds. Jay M. Heubert and Robert M. Hauser. Washington, DC: National Academy Press, 1999), in response to a Congressional charge “to recommend appropriate methods, practices, and safeguards to assure that existing and new tests ... are not used in a discriminatory manner or inappropriately for student promotion, tracking, or graduation, and existing and new tests adequately assess student reading and mathematics comprehension in the form most likely to yield accurate information regarding student achievement of reading and mathematics skills” (P.L. 105-78, Sec. 309). The NRC panel was a diverse group of 15 scholars from all across the country. My comments on the Chicago study are informed by the NRC study, but they are mine alone. Neither the NRC nor my fellow committee members bear any responsibility for my opinions.

When used appropriately, high-stakes tests can help promote student learning and equal opportunity in the classroom by defining standards of student achievement and by helping school officials identify areas in which students need additional or different instruction. When used inappropriately, high-stakes tests can undermine the quality of education and reduce opportunities for some students, especially if results are misinterpreted or misused, or students are relegated to a low-quality educational experience as a result of their scores.

One of the NRC's strongest recommendations was that "Accountability for educational outcomes should be a shared responsibility of states, school districts, public officials, educators, parents, and students. High standards cannot be established and maintained merely by imposing them on students" (p. 5). Chicago has tried to raise achievement largely by imposing arbitrary standards on students. It is time for Chicago's educational leaders to become accountable, to recognize the failure of their "get-tough" policies, and to turn in new directions.

A handwritten signature in black ink, reading "Robert M. Hauser". The signature is written in a cursive style with a large initial "R" and "H".

Robert M. Hauser