DOCUMENT RESUME

ED 470 555 UD 035 348

TITLE	The Minnesota Basic Skills Test: Performance Gaps on the Reading and Mathematics Tests from 1996 to 2001 by Gender, Ethnicity, Limited English Proficiency, Individual Education Plans, and Socio-Economic Status.
INSTITUTION	Minnesota Univ., Minneapolis. Office of Educational Accountability.
PUB DATE	2002-07-00
NOTE	54p.
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Descriptive (141)
EDRS PRICE	EDRS Price MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Academic Standards; *Basic Skills; Disabilities; *Grade 8; Junior High School Students; Junior High Schools; Limited English Speaking; Mathematics Achievement; Minority Group Children; Racial Differences; Reading Achievement; Sex Differences; Socioeconomic Status; Standardized Tests; *State Standards; Tables (Data)
IDENTIFIERS	*Achievement Gap; Minnesota

ABSTRACT

This study examined emerging educational trends in Minnesota following the implementation of the Minnesota Basic Skills Tests. Data consisted of demographic and test score information for all 8th grade students taking the regular administration of the mathematics and reading Basic Skills Tests from 1996-01. Students' performance in reading appears to be improving over time, though this increase is leveling off. Mathematics performance has shown no clear pattern over the study period, though there are hints at improvement. Gender difference in mathematics have consistently decreased over time. Gender differences in reading favor females and are not decreasing. The relative percentage of white students is decreasing over time, as Minnesota becomes a more diverse state. Test performance differences between black and white students are large, especially in mathematics. Students with limited English proficiency (LEP) and students with individual education plans (IEPs) have lower test scores than their typical peers, though the performance gap is decreasing for LEP students. Students with both LEP and IEPs perform at extremely low levels. Students from low-income families have lower test scores than their counterparts who are not low income. (Contains 18 figures, 31 tables, and 15 references.) (SM)

THE MINNESOTA BASIC SKILLS TEST: PERFORMANCE GAPS ON THE READING AND MATHEMATICS TESTS FROM 1996 TO 2001, BY GENDER, ETHNICITY, LIMITED ENGLISH PROFICIENCY, INDIVIDUAL EDUCATION PLANS, AND SOCIO-ECONOMIC STATUS

July 2002

Prepared by

The Office of Educational Accountability
College of Education and Human Development
University of Minnesota

The Minnesota Basic Skills Test: Performance Gaps for 1996 to 2001 on the Reading and Mathematics Tests, by Gender, Ethnicity, Limited English Proficiency, Individual Education Plans, and Socio-Economic Status

July 2002

Prepared by

The Office of Educational Accountability
College of Education and Human Development
University of Minnesota

OEA Staff:
Mark L. Davison, Ph.D., *Director*
Ernest C. Davenport, Ph.D., *Associate Professor* (primary author)
Nohoon Kwak, Ph.D., *Statistician*
Kristin Peterson, *Research Associate*
Margaret L. Irish, *Editor*

Chi-keung Chan, *Graduate Assistant*
Jiyoung Choi, *Graduate Assistant*
Kamil Guven, *Graduate Assistant*
Jeffrey Harring, *Graduate Assistant*
Yun Jung Kang, *Graduate Assistant*
Yi-ChenWu, *Graduate Assistant*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Given that the Minnesota *Basic Skills Tests (BSTs)* have been in place for over half a decade, it is reasonable to desire information on emerging trends. This manuscript begins with an explanation of the philosophical foundations of Minnesota's graduation rule. It then provides a historical look at the math and reading portions of the *BST* and identifies several trends relating to time, gender, ethnicity, limited English proficiency status (LEP), individual education plan status (IEP), and whether the student was eligible for free or reduced-price lunch (a surrogate for socio-economic status [SES]). Data were obtained for the reading and mathematics portions of the *BST* for 8^{th} grade students taking the regular administration of the test for the years 1996–01. The basic statistic used was the effect size, an index that gives the distance between the means for two groups, in terms of their standard deviation. Thus, each comparison is on a standard scale. Finally, the results are put into context relative to national data and President Bush's educational initiative, *No Child Left Behind.*

Throughout this paper, attention is focused on the trends themselves rather than on explanatory factors (e.g., the steady increase in the number of students in Minnesota who have limited English proficiency, rather than the possible causes of this increase). In the current "accept-no-excuses" environment of school accountability, the focus is on closing gaps, not on possible explanations for them. In line with this approach, this paper was designed to describe current discrepancies. The main results are as follows.

Data for the first two years of testing, 1996 and 1997, may not be comparable to data for the final four years (1998–01) due to a change in students taking the exam. The *Basic Skills Test (BST)* was not mandatory for the first two years. Students could use their scores on other tests to satisfy the high school graduation requirement. However, after 1997, the *BST* became the only way to satisfy the requirement.

Students' performance in reading appears to be improving over time. This increase in scores, however, seems to be leveling off.

Students' performance in mathematics has shown no clear pattern over the period studied. However, the difference between the last two years' scores (1999–01) hints at an improvement that was confirmed with the 2002 results (announced as this report was being completed).

Gender differences in mathematics favor males. However, this difference was small initially, has consistently decreased over time, and has now almost vanished. This finding is consistent with the goal of closing achievement gaps as outlined in President Bush's *No Child Left Behind* act.

Gender differences in the *BST* in reading are larger than gender differences in mathematics, and these differences favor females. Unlike the differences in math, gender differences in reading are not decreasing over time: the difference in 2001 is the same as the difference in 1996. However, the difference in reading scores is still small enough to be classified as a "small" difference (an effect size of 0.20 or less).

In national tests given by the National Assessment of Educational Progress (NAEP), gender differences for reading also favored females. As with the *BSTs*, NAEP results for mathematics favored males. Similarly, the difference on the NAEP reading assessment was larger than the difference on the NAEP mathematics assessment. Finally, the absolute differences were larger for NAEP, indicating a larger gender difference than for the *BSTs* in both math and reading.

Consistent with census findings on total population, the relative percentage of White students is decreasing over time, as Minnesota becomes a more diverse state. The percentage of each minority group is rising, except for the Native American population, which retained virtually the same percentage from 1990–00.

Test performance (highest to lowest) for students of the various ethnic groups is as follows: Whites, Asians, Native Americans, Hispanics, and Blacks. The performance of the Asian group places them approximately halfway between the White group and the Native American group. The performance of the Native American and Hispanic groups are relatively close. Finally, the Black group is by itself at the bottom. The difference in performance between Black and White students is huge, especially in mathematics.

In mathematics, there is a trend toward equity for Blacks and Asians relative to White students. Hispanics and Native Americans show a slight trend away from ethnic equity. In reading, there is a trend away from equity for Native Americans and Hispanics, as compared against White students. Blacks and Asians have a slight trend toward ethnic equity.

NAEP data shows a decrease in national performance differences between Black and White students from 1970 until approximately 1989, when the gap began widening again.

Effect size results from NAEP for 13 year olds over the last decade show a widening of the national racial gap between Whites and Native Americans, as well as between Whites and Hispanics. The Black/White gap is closing, but it is the largest of the gaps. Finally, the Asian/White gap is closing; however, the Asian group started with scores *higher* than those for Whites.

Students with limited English proficiency (LEP) do less well in both reading and mathematics than students who are not classified as LEP. As expected, these students score relatively lower in reading than in mathematics. However, the difference between the gaps for reading and mathematics is small.

Minnesota is headed in the right direction, decreasing the performance gap between LEP-eligible and non-LEP-eligible students in both mathematics and reading. These gains appear to be real performance gains for reading. In math, however, part of the gain may be due to changes in classifying students. In any case, the size of the disparity and the rate of decrease will make this a long-term effort.

Students with individual education plans (IEPs) score less well in both reading and mathematics than students without IEPs. The discrepancies are almost twice as great as those Cohen (1988) defined for "large" effects (the threshold for a large effect is 0.80).

In general, the discrepancies for IEP-eligible versus non-IEP-eligible students are larger than that for LEP-eligible versus non-LEP-eligible students.

Over the last four years of testing (1998–01), the trend for IEP-eligible versus non-IEP-eligible students in mathematics is in the correct direction, although the ending difference is still huge. The trend for reading, however, shows ever-widening gaps between IEP-eligible and non-IEP-eligible students.

Minnesota's academically neediest students, those with both limited English proficiency *and* individual education plans, are performing at extremely low levels. The effect size difference between students classified with both LEP and IEP compared to students without either of these classifications is close to 2.5 standard deviation units for both mathematics and reading.

The trend toward equity between students with both IEP and LEP status compared to those without either classification is in the right direction for math; however, the gap is huge. The 2001 effect size is 2.4. For reading, the trend is not as favorable. There is a consistent rise in the effect size difference (i.e., away from equity), reaching almost 3 standard deviation units in 2000 before falling to 2.7 in 2001. It is important to keep in mind that the small sample size for students with both IEP and LEP status limits generalization of these results.

Due to the interrelatedness of ethnicity with both IEP and LEP status, another analysis was conducted to compare ethnic differences for students without either of these classifications. The results found a decrease in the effect size difference between White versus Asian and Hispanic students. In contrast, the effect sizes for Blacks and Native Americans relative to Whites were larger. This increase in effect size was, in part, an artifact of the comparison procedure.

Students eligible for free or reduced-price lunch score lower than their counterparts who are not so classified. The difference in effect size is over 0.80—a large effect. Although the discrepancies for reading and mathematics are similar in magnitude, they are a little larger for mathematics.

The Minnesota data suggests that the socio-economic status (SES) gap appears to be widening for performance on both math and reading assessments. The trend and size of the effects are consistent with those apparent in the NAEP data. There, the effect sizes are also close to 0.80, and for mathematics (based on two years of data), the gap is also widening.

Both Minnesota and the nation have much work to do to close achievement gaps.

INTRODUCTION

Given that the Minnesota *Basic Skills Tests* have been in place for over half a decade, it is reasonable to want information on emerging trends. Identification of emerging trends would be useful as Minnesota works to continuously improve its education system for all students. In addition to the usefulness of these emerging trends in state accountability efforts, there is now an external impetus for identifying emerging trends: the new educational thrust of President Bush's *No Child Left Behind* initiative (Bush, 2001). This initiative was signed into law on January 8, 2002; it aims to reward education systems for closing achievement gaps between diverse subgroups. In contrast, money will be withheld from states not closing these gaps.

Minnesota's current push toward a performance-based educational accountability system has its roots in the Outcomes-Based Education (OBE) movement of the 1980s. Minnesota created a performance-based system whereby desired outcomes could be directly specified in terms of specific performance requirements. An endorsement of this effort was codified into law in 1992, when the Minnesota Legislature approved a standards-based *Graduation Rule*, which included two distinct components. The first component is the *Basic Standards*, representing the minimum skills required of all students for high school graduation. The second component contains *Preparatory Standards* for grades K–8, and *High Standards* (also required for high school graduation) for grades 9–12 (Davison, Erickson, Davenport, Kwak, Irish, Bielinski, Danielson, Kim, Seo, Smith, and Wick, 1999).

After passage of the legislation, the Department of Children, Families & Learning began fulfilling the requirements set forth in the legislation, while remaining true to the philosophical foundations of performance-based assessment. The Minnesota *Basic Skills Tests (BSTs)* in reading, mathematics, and writing were developed to assess the minimum standards necessary for all high school graduates (with some exceptions, e.g., certain special education students). A passing score on each of the tests is needed for high school graduation. In contrast to the measures of basic standards, "performance packages" were designed to measure each student's attainment of the Preparatory and High Standards. These performance packages consist of tasks that students must complete satisfactorily in order to meet requirements in various content areas.

The *Basic Skills Tests* in mathematics and reading have been given at least once a year since 1996. The tests are taken by most 8^{th} graders as well as by students in other grades. Some students in grades 9–12 take the test for a variety of reasons (for example, having moved into the state after grade 8, or not having passed on their first attempt at the *BSTs*). There have also been special administrations of the tests in the summer, for students who have not passed the tests; and immediately before the end of the school year, to give high school seniors who have yet to pass some portion of the test one last chance to graduate with their class.

The following historical look at the math and reading portions of the *Basic Skills Tests* provides performance information for different subgroups of students. This information leads to the identification of several trends that should be of interest to administrators, teachers, parents, and students. Of special interest are the achievement information relative to different "at-risk" groups, and the corresponding trend information on the shrinking or widening of achievement gaps.

The paper begins with the methodology used for the study. The findings section summarizes the trends in math and reading for several demographic characteristics, namely, ethnicity, gender, limited English proficiency (LEP) status, individual educational plan (IEP) status, and socio-economic status (SES). When possible, the results are placed in context, with national trends, as reflected in test scores from the National Assessment of Educational Progress (NAEP).

It is worth emphasizing that this paper addresses only the trends; it does not explore the reasons for them since the current "accept-no-excuses" environment of school accountability focuses on closing gaps rather than on possible explanations for them.

METHODOLOGY

Data

Data for this study consist of demographic and test score information for all 8th grade students taking the regular administration of the *Basic Skills Tests* in mathematics and reading over the years 1996–01. The analyses were restricted to 8th graders in an attempt to make the data comparable year-to-year. For each year since 1996, students other than 8th graders were tested. Some of these students are repeat test takers. Given that the repeat test takers are usually students who have not passed the test, the varying numbers of these students will probably influence average estimates of performance. To rid the data of this possible confound, the present study was limited to data from 8th graders. Scores for students participating in the special testing sessions (summer and April) are excluded for similar reasons. It is our belief that these test-takers are fundamentally different from the group of 8th graders who take the test during the regular administration.

Finally, with respect to the first two years of data, caution is required in interpreting the results. During the first two years (1996 and 1997), the *Basic Skills Test* was not compulsory; students had the option of using their scores on other tests to meet the high school graduation standard.

Although an effort has been made to use comparable groups of students from one year to the next, it should still be stressed that the basic design is cross-sectional (i.e., a different group of students is assessed each year). From one year to the next, therefore, the students may be fundamentally different in experiences, background, and in the relative composition of various subgroups. Thus, some of the differences in performance from year to year may be an artifact of the study subjects (students) selected, rather than a statement about the educational system.

Data Analysis

This study uses a descriptive methodology, consisting of trends in mean differences between various student subgroups. The means are based on "percent correct" scores. For math, the percentage is based on 68 items. There are 40 items on the reading test.

The main statistical index used is the effect size (Cohen, 1988). Effect sizes give the relative difference in means for different groups, based on standard deviation units. To illustrate, if two groups have the same standard deviation (10 points) for a measure of performance, and Group A has a mean of 80 while Group B has a mean of 85, then the effect size is (85 – 80) / 10 = 0.50. Thus, the group means are half a standard deviation unit apart. This interpretation does not depend on how the variables were scored originally, so use of effect size allows all comparisons to be reduced to a single, standardized scale that represents the *relative difference* between the two groups in standard deviation units.

Moreover, following the suggestions of Cohen (1988), it is possible to label effect sizes by degree. For mean differences, an effect size of 0.20 is small; 0.50 is medium; and 0.80 is large (Cohen, 1988). Moreover, if we assume that the scores from our groups follow a "bell curve," Table 1 gives the percentage of overlap for two groups of students with differing performance, as represented by the specific effect size differences. Table 1 also gives the percentage of scores for students in the low scoring group, whose scores equal or exceed the mean for students in the higher scoring group for various effect size values. Figure 1

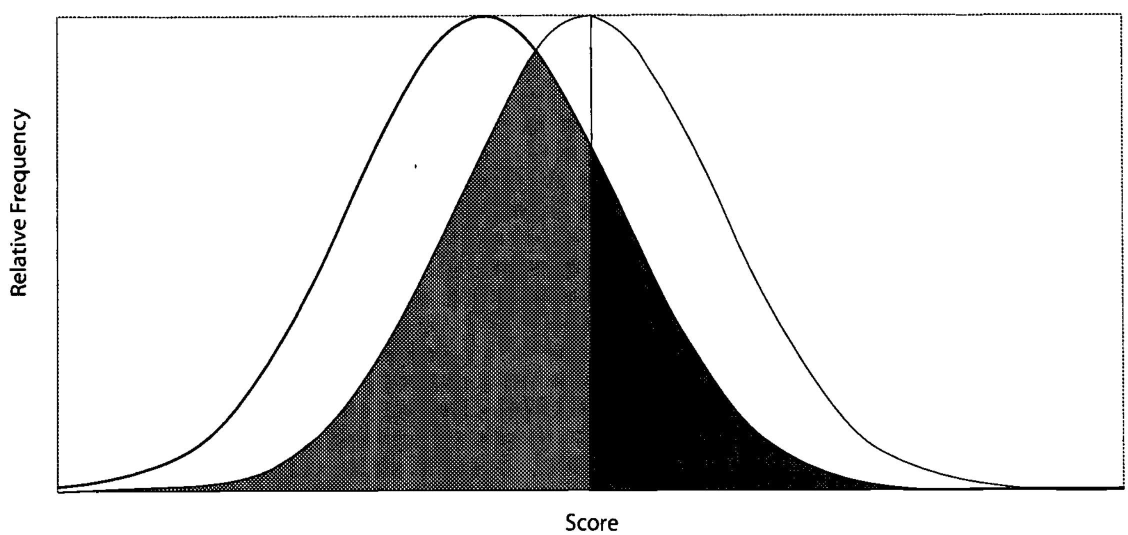
Table 1. Effect Size Statistics

Effect Size	Size of Effect*	% Overlap of Distributions	% in Lower Group to Exceed the Mean for the Higher Group
0.2	Small	85.30%	42.10%
0.5	Medium	67.00%	30.90%
0.8	Large	52.60%	21.20%
1		44.60%	15.90%
1.5		29.30%	6.70%
2		18.90%	2.30%
2.5		11.80%	0.60%

** Assuming Normally Distributed Scores for Both Groups Cohen (1988)*

illustrates the difference for a large effect size, 0.80. Note the overlap in the two distributions—the shaded area in Figure 1. It comprises about half of the total area for both distributions. The darker shaded region represents scores for students falling in the lower-scoring distribution that exceed the mean score for the group of higher scoring students. If both distributions have the same mean (and are symmetric), one would expect about half of the scores to be at or above the mean. However, if the two groups of students are 0.80

Figure 1
Distributional Overlap: Large Effect



standard deviation units apart, only a fifth of the students in the lower group would be expected to have scores at or above the mean for the higher group.

FINDINGS

Annual

Table 2 shows the mean (Mean), standard deviation (SD), sample size (N), and percentage of the total sample (% of Sample) for the 8th grade students who took the regular administration of the math and reading tests for each of the years 1996 to 2001. The variable of interest is the percentage correct.

For mathematics, the number of students (N) was 58,465 the first year, 51,923 the second year, and over 64,000 each of the last four years. The number of test takers in reading shows a comparable pattern. The discrepancy in the number of subjects in 1996 and 1997, as opposed to the final four years (1998–01), is due to the fact that the test was not mandatory for the first two years. In 1996 and 1997, it was possible to use scores on other tests to satisfy the testing requirement. After the first two years, the *BST* became the only test that could satisfy the testing requirement. It is worth noting that this may affect some conclusions in the subsequent analyses, since the population of students tested in the first two years may be fundamentally different from the population tested later.

Table 2 shows that the mean score for each year in both mathematics and reading is over the 70% correct passing standard set for students whose graduation year was 2000. After 1997, the mean for both scores is over the 75% mark set for subsequent graduating classes (2001 and beyond). However, the mean score for reading in 1997 falls a little short of this standard, at only 74.27%.

Table 2. Descriptive Statistics for Mathematics and Reading *Basic Skills Tests* for All Minnesota Eighth Graders (1996–01)

	Mathematics				Reading			
Year	N	% of Sample	Mean	SD	N	% of Sample	Mean	SD
1996	58,465	**15.79%**	78.85	15.58	56,516	15.39%	71.79	18.09
1997	51,923	**14.02%**	79.49	15.69	50,380	13.72%	74.27	17.67
1998	64,396	**17.39%**	79.03	17.95	64,403	17.54%	77.96	17.89
1999	65,361	**17.65%**	79.06	18.07	65,404	17.81%	80.73	16.95
2000	65,913	**17.80%**	79.00	16.69	65,976	17.97%	82.53	16.06
2001	64,311	**17.36%**	80.01	16.47	64,509	17.57%	82.93	17.04
Total	*370,369*	*100.00%*	*79.24*	*16.83*	*367,188*	*100.00%*	*78.69*	*17.26*

Note that although the original passing standard was given in terms of percentage correct, these values are no longer used to make this determination. Scaled scores are now used to provide more comparability in actual performance from one year to the next. The passing value for scaled scores do not always correspond perfectly to the 70 and 75 percent values; however, the fluctuation has been at most one item and there has been only one occurrence for the two tests and multiple administrations per year where use of scale scores changed the percentage correct needed to pass. Specifically, the passing standard for math in 2000 based on the scaled score was one item less than 75%.

Percentage correct was used in this paper because it is easy to interpret; the intent of this manuscript is to ascertain performance level, rather than merely whether the student passed or not. Initially, all analyses were conducted using the percentage correct as well as whether the student passed. The latter analyses were dropped from this manuscript, because a pass/no pass decision effectively eliminates most of the information in the score. Further distinctions can be made based on the fact that all who pass are not the same, just as all who fail are not the same. Given our desire to use an index that contained full information, either the raw score, the percentage correct, or the scaled scores are essentially equivalent. Percentage correct is the simplest and was chosen for that reason. At times, we relate the percentage correct to the passing standard (scale scores).

The most interesting finding in Table 2 is the relatively constant math scores, as compared to the consistently rising reading scores over time. The range in scores for mathematics is only 1.16% (78.85 – 80.01). In contrast, the mean reading score changes by 11.14% (71.79 – 82.93). The change in reading test scores over the six years is almost 10 times that of the mathematics test scores for the same time period. Figure 2 gives a picture of these relationships. While the line for mathematics is almost flat, the line for reading shows a clear, consistent increase. Note, too, that the average percentage correct for reading started lower than the average percentage correct for mathematics. The reading score, however, subsequently passed the math score.

Figure 2
Annual Average Percentage Correct for Mathematics and Reading (1996–01)

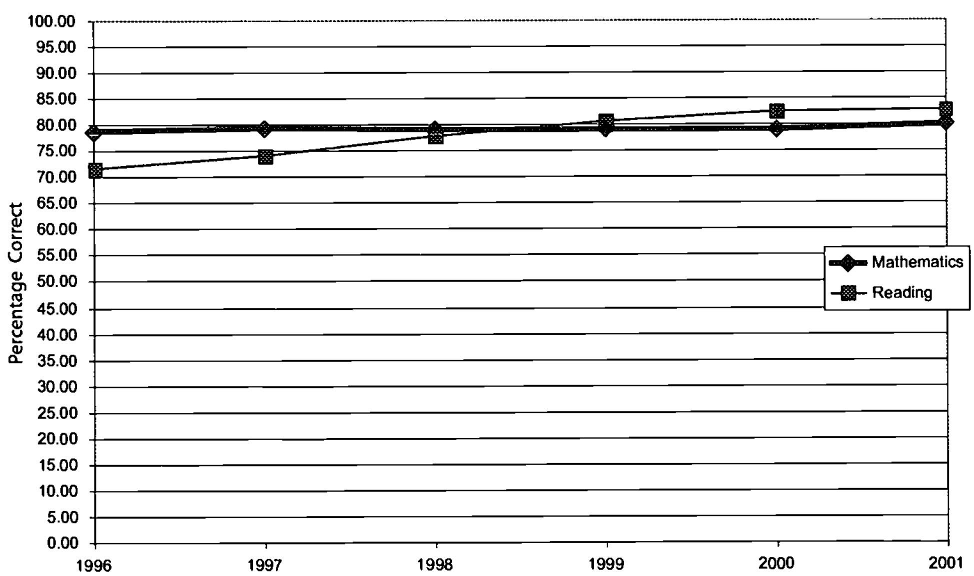


Table 3 (p. 9) summarizes the effect size difference between years for both math and reading. Negative values indicate a drop in performance over time. For example, the first value in the table, 0.04, shows an increase from 1996 to 1997. The –0.03 value (one step to the right and one step down) indicates a score decrease from 1997 to 1998. For reading, there is a 0.14 effect size gain between 1996 and 1997, a 0.21 gain between 1997 and 1998, a 0.16

Table 3. Effect Size Changes for Mathematics and Reading (1996–01)

	Math					Reading				
	1997	1998	1999	2000	2001	1997	1998	1999	2000	2001
1996	0.04	0.01	0.01	0.01	0.07	0.14	0.36	0.52	0.62	0.65
1997		-0.03	-0.03	-0.03	0.03		0.21	0.37	0.48	0.50
1998			0.00	0.00	0.06			0.16	0.26	0.29
1999				0.00	0.06				0.10	0.13
2000					0.06					0.02

** Negative values denote a decrement in performance from one year to the next.*

gain between 1998 and 1999, a 0.10 gain between 1999 and 2000, and a 0.02 gain between 2000 and 2001: that is, there was progress each year. From 1996–01, the cumulative effect size change in the reading score is just over half of a standard deviation unit, 0.65 (a medium effect). Finally, we see that the gain in performance from year to year is leveling off. The magnitude of the gain rose in the first two years, from 0.14 to 0.21, and then began to decrease with each succeeding year, from 0.16 to 0.10, with a final gain of 0.02 in 2001. This consistent decrease in performance gains suggests that reading performance may be approaching a ceiling. Thus, future gains in reading performance may be harder to achieve. However, the increase in reading scores is encouraging; it means that students are better prepared for the tests. It is our belief that schools' and districts' efforts to improve the basic academic skills of students in Minnesota are proving effective for reading (see Schleisman, Peterson, and Davison, 2000).

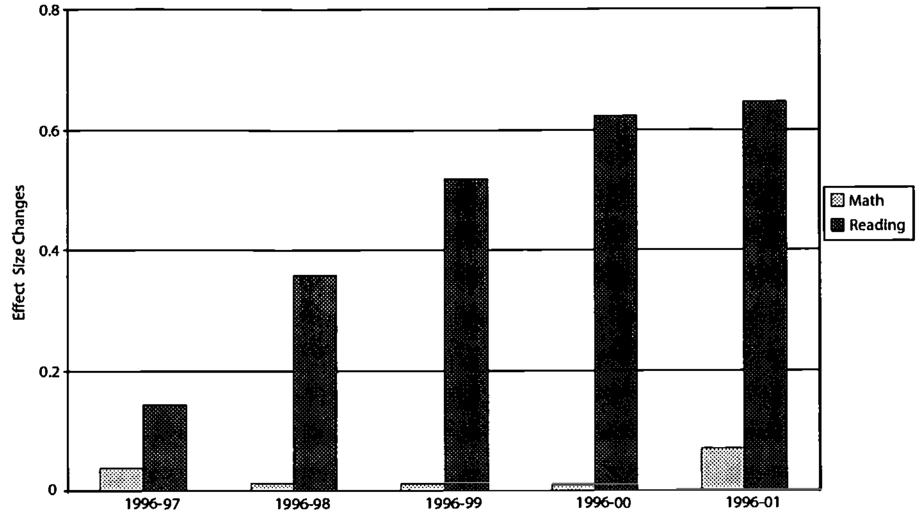
The performance change in math is very small. The effect size gain over the six years (1996–01) is only 0.07, and most of this change occurred in the final year, 2001. The good news is that the resulting change is positive. There is little discernible trend, however, from 1996–00: the year-to-year change is sometimes negative, sometimes positive, and sometimes shows no change at all. It is especially interesting that the average score change from 2000–01 is the most positive gain for mathematics. The raw score change is over one full percentage point, which corresponds to an effect size of 0.06. This result may be an aberration—or a trend that is just beginning. The results for the 2002 assessment will help to shed light on the correct interpretation of this final positive gain—that is, whether it is an aberration or a nascent trend.[1] Figure 3 (p. 10) illustrates the difference in cumulative effect size for performance on the mathematics and reading tests from 1996–01. The gain in reading is nine times the gain in mathematics (0.65 versus 0.07).

Schleisman, et al. (2000) detail some of the extraordinary remediation efforts initiated by schools and districts in Minnesota as a direct result of implementation of the *Basic Skills Tests*. Some of these efforts include: a) basic skills classes offered during the regular school day, b) remediation within regular classes—students are placed into these classes based on ability, c) a focus on reading and math across the curriculum, d) summer school programs, e) after-school and Saturday programs, f) resource rooms, and g) study packets. With the exception of a report on the effectiveness of summer school interventions (Davison, Schleisman, Koeppen, Wu, and Kwak, 2001), there has been no evaluation of the effect of these extra efforts. However, the results reported above seem to suggest that the combined

NOTES

[1] An increase in the math score for 2002 was confirmed in data released while this report was being completed.

Figure 3
Cumulative Effect Size Changes for Mathematics and Reading (1996–01)



result of these efforts, along with regular schooling, seems positive for reading, but less effective for mathematics.

Gender

Because of equity issues, it is important to know if there are differential gender trends that are observable in the data over the past six years. Table 4 (p. 11) gives descriptive statistics for reading and math for males and females. Table 5 (p. 12) shows the corresponding effect size information. The mean scores for both reading and math are over the 70% correct mark, set as passing for students with a graduation year of 2000. This is true for males and females, both overall and for each year. Note, however, how close the mean reading score for males is to this mark in 1996 (70.3%). Note, too, that the mean reading score for males in 1997 misses the 75% correct mark set for the 2001 graduates, the actual value being 72.4%. All of the other annual results show mean scores in excess of 75% for both males and females on both reading and math.

Looking at Table 5 , we find that the overall effect size difference for math, although small, favors males. The effect size is 0.04. In contrast, the reading difference is four times as large as for math, 0.16, and favors females. While the effect size for reading is larger, it is still small enough to be classified as a small effect (Cohen, 1988).

Another valid inquiry is the change in effect size by gender over time. Table 5 shows the effect sizes for males versus females on the math and reading tests over the years 1996–01. Although our previous results show that the overall scores for math change little over time, the discrepancy between males and females is shrinking consistently. In 1996, the difference was 0.07 standard deviation units in favor of males. This difference drops each year until it stands at 0.01 in 2001. The results for reading are different: the discrepancy is larger and favors females. The discrepancy in performance for reading increased from 1996–97 (from an effect size of 0.17 to 0.22), and then decreased in 1998–99 before immediately increasing again in 2000–01. The net result of all of the changes is that after six years of testing, the effect size difference between males and females for reading is essentially the same, 0.17 standard deviation units, as it was initially.

Table 4. Means and Standard Deviations, by Gender (1996–01)

		Mathematics				Reading			
Year	Gender	N	%	Mean	SD	N	%	Mean	SD
1996	Male	30,455	52.09%	79.39	16.03	29,494	52.19%	70.31	18.73
	Female	28,010	47.91%	78.25	15.06	27,022	47.81%	73.39	17.22
		58,465	100.00%	78.85	15.57	56,516	100.00%	71.79	18.02
1997	Male	27,545	53.05%	79.92	16.21	26,495	52.59%	72.42	18.45
	Female	24,378	46.95%	78.99	15.08	23,885	47.41%	76.33	16.52
		51,923	100.00%	79.49	15.69	50,380	100.00%	74.27	17.56
1998	Male	32,362	50.97%	79.77	17.93	32,412	51.00%	76.98	18.45
	Female	31,131	49.03%	78.76	17.57	31,144	49.00%	79.41	16.88
		63,493	100.00%	79.27	17.76	63,556	100.00%	78.17	17.70
1999	Male	33,450	51.20%	79.44	18.42	33,442	51.15%	79.97	17.73
	Female	31,878	48.80%	78.67	17.68	31,932	48.85%	81.55	16.04
		65,328	100.00%	79.07	18.06	65,374	100.00%	80.74	16.93
2000	Male	33,876	51.41%	79.15	16.94	33,878	51.36%	81.23	17.15
	Female	32,017	48.59%	78.86	16.42	32,081	48.64%	83.92	14.69
		65,893	100.00%	79.01	16.69	65,959	100.00%	82.54	16.00
2001	Male	33,171	51.58%	80.12	16.85	33,245	51.54%	81.53	17.93
	Female	31,139	48.42%	79.89	16.05	31,263	48.46%	84.41	15.90
		64,310	100.00%	80.01	16.47	64,508	100.00%	82.93	16.98
Overall	Male	190,859	51.67%	79.62	17.13	188,966	51.59%	77.39	18.54
	Female	178,553	48.33%	78.91	16.42	177,327	48.41%	80.16	16.65
		369,412	100.00%	79.28	16.79	366,293	100.00%	78.73	17.65

Figure 4 (p. 12) shows two bars for each year. The bars represent the difference between males and females in standard deviation units for the given content areas (math/reading) for a given year. The left-hand bar of each pair is darker and represents mathematics—it shows the male advantage. The right-hand bar of each pair is light gray and represents reading—this bar shows the female advantage. Figure 4 suggests that gender equity is becoming a reality for mathematics achievement in Minnesota in the *Basic Skills Test*. This is not true for reading. Although the difference decreased in 1998 and 1999, the difference in 2000 and 2001 was approximately the same as in 1996. We stress that the difference is, by definition, a small effect. However, it still would be more desirable if both the math and reading differences were shrinking.

For several subgroup comparisons, national data exist that make it possible to place the *BST* results in the national context. The national data are from NAEP, the National Assessment of Educational Progress. NAEP is often called the *Nation's Report Card*. It was created by the National Center for Educational Statistics (NCES), which is charged with

Table 5. Effect Size Differences for Gender (1996-01)

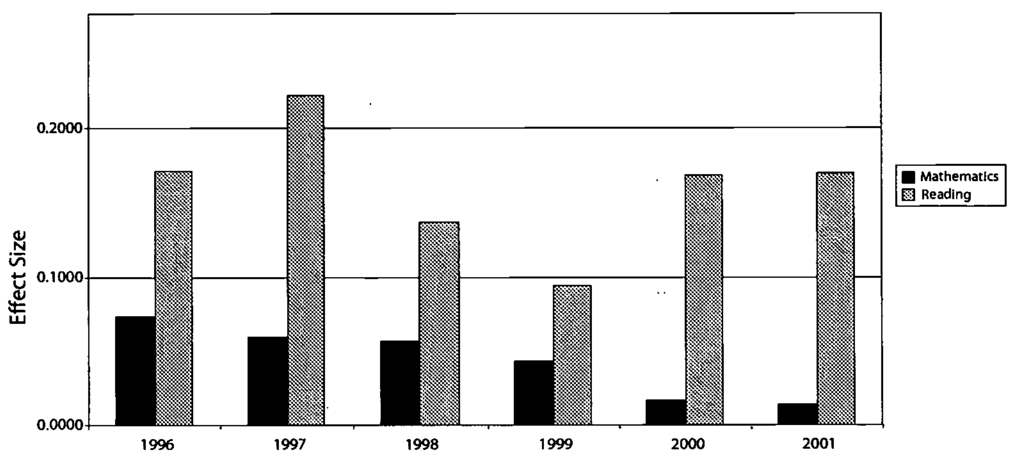
Year	Mathematics	Reading
1996	0.07	-0.17
1997	0.06	-0.22
1998	0.06	-0.14
1999	0.04	-0.09
2000	0.02	-0.17
2001	0.01	-0.17
Overall	0.04	-0.16

Positive values indicate higher scores for males. Negative values indicate higher scores for females.

collecting data on the nation's educational progress. On a regular basis, NAEP assesses the nation's youth in several learning areas, including mathematics and reading, and its random sample of 13 year olds makes a natural comparison group for Minnesota 8th graders who have taken the *BST*.

NAEP's mathematics tests have five strands: 1) number sense, properties, and operations; 2) measurement; 3) geometry and spatial sense; 4) data analysis, statistics, and probability; and 5) algebra and functions. The question format is a combination of multiple choice and constructed response. The reading tests are composed of complete texts from typical grade-appropriate sources. There are two types of texts representing different purposes for reading: 1) reading for literary experience and 2) reading for information. Again, there are both multiple choice and constructed response questions. All of the NAEP data presented in this report are obtained from a data-gathering tool (available at http://nces.ed.gov/naep3/naepdata/search.asp). The latest mathematics data were for 2000; the latest reading data were for 1998.

Figure 4
Effect Sizes in Mathematics and Reading: Comparison of the Differences in Performance of Boys and Girls (1996–01)



Note: the bar heights represent the size of the difference between boys' and girls' performance in mathematics and reading. The differences between boys' and girls' performance constitute only a small effect (Cohen, 1988) for both reading and math.

The gender differences in performance for the *BST* are mirrored in the NAEP data. NAEP's 2000 math test shows an effect for 13 year olds of 0.08 in favor of males. The 1998 reading results show an effect size of 0.39 in favor of females. The means, sample sizes, and standard deviations necessary to calculate the effect sizes are found in NCES (2002). For NAEP, the discrepancies are in the same direction as for the *BST* (i.e., math favors males while reading favors females). Moreover, the reading difference is again larger than the mathematics difference. Table 6 summarizes the *BST* and NAEP results.

The gender difference is less for the *BST* than for NAEP, and the gender effect in reading for NAEP is twice as large as for the *BST* (0.39 versus 0.17). In

Table 6. Effect Size Comparisons for *BST* and NAEP Data

	2001 BST Minnesota	NAEP National
Reading	0.17	0.39
Math	0.01	0.08

The NAEP Reading data is from 1998 and the math data is from 2000.

fact, the NAEP difference is no longer a small effect; it is between a small and medium effect (Cohen, 1988). Likewise, the NAEP math effect favoring males is larger than for the *BST*—a difference of 8 to 1 (0.08 vs. 0.01). These data suggest that Minnesota may have difficulty closing its gender gap in reading. Furthermore, the national gaps are larger than the Minnesota gaps for both math and reading. The larger national gaps imply that Minnesota is not alone in its need to close these gaps, and that Minnesota has less of a problem than does the nation as a whole.

Ethnicity

Much has been written about ethnic differences as they relate to performance on the *BSTs*. Two examples are reports by The Roy Wilkins Center for Human Relations and Social Justice at the University of Minnesota. The first, entitled *Summary of Results: An Analysis of the 1996 Minnesota Basic Skills Test Scores*, was published in March 1997. The second, *Racial Disparities in Minnesota* Basic Standards Test *Scores, 1996–2000*, was published in October 2000. The first paper examined score differences by gender; ethnicity, limited English proficiency, and individual education plan status. However, the bulk of that work presented models to "answer questions about the determinants of performance and the factors contributing to the wide racial gaps in mathematics and reading scores." The goals of the second paper were to provide an update to the first and to note trends that related to the different ethnic groups. In contrast, the goal of this report is to provide a description of trends where one of the components is ethnicity.

Table 7 (p. 14) gives descriptive statistics for the reading and math tests by ethnic group. Overall, 1.75% of the students taking the tests were Native Americans. This figure is 4.45% for Asian students, 2.02% for Hispanics, 4.73% for Blacks, and 87.05% for Whites. The annual numbers for the minority groups are rising. While Whites represented over 89.5% of those tested in 1996, this value had decreased to 85% in 2001. Part of this decrease could reflect the fact that the tests were not mandatory the first two years and, therefore, relatively more majority students may have been tested during that time (1996 and 1997). However, another factor is demographics.

Census data show that the percentage of minorities is rising in Minnesota. Table 8 (p. 15) presents the census results for 1990 and 2000. During that time the percentage of Minnesotans who were White went from 93.7% to 89.6%. Table 8 also shows that the only minority group that did not have a relative increase in numbers was Native Americans. Their representation in the population remained almost unchanged over the ten years from 1990 to 2000 (1.13% versus 1.12%). Note that the number of Native Americans did increase, but so did the number for all ethnic groups. Unlike the other ethnic minority groups, the relative increase for Native Americans did not exceed the relative increase for the total population.

The mathematics achievement results in Table 7 show that only the White and Asian groups had a mean score above the 70% passing mark set for 8th grade students in 1996. For reading, the only group with a mean score higher than the cut score was the White group. When the passing mark was raised to 75%, only the White group's mean rose above passing in either mathematics or reading between 1997 and 1999. In 2000 and 2001, the Asian group joined the White group in achieving means above 75% on both the mathematics and reading tests. No other ethnic group had means above the passing mark during the six years (1996–01).

Table 7. Descriptive Statistics for the Scores, by Ethnicity (1996-01)

		Mathematics				Reading			
		N	%	Mean	SD	N	%	Mean	SD
1996	Native American	934	1.64%	66.89	17.83	898	1.63%	61.47	19.45
	Asian	2,050	3.61%	72.59	17.40	1,990	3.62%	64.27	20.29
	Hispanic	864	1.52%	66.64	17.91	821	1.49%	60.16	19.46
	Black	2,103	3.70%	58.63	18.66	2,009	3.65%	53.54	19.94
	White	50,908	89.53%	80.62	14.14	49,273	89.60%	73.44	17.07
		56,859	**100.00%**	**79.08**	**14.59**	**54,991**	**100.00%**	**71.99**	**17.38**
1997	Native American	682	1.35%	69.13	17.13	603	1.23%	63.52	19.03
	Asian	2,272	4.51%	73.01	17.90	2,264	4.63%	64.75	20.79
	Hispanic	856	1.70%	66.65	18.01	832	1.70%	62.74	19.33
	Black	2,374	4.72%	58.98	17.75	2,378	4.86%	56.32	19.64
	White	44,151	87.71%	81.59	14.05	42,836	87.58%	76.41	16.21
		50,335	**100.00%**	**79.72**	**14.56**	**48,913**	**100.00%**	**74.50**	**16.73**
1998	Native American	1,192	1.87%	65.05	21.02	1,185	1.86%	65.03	20.08
	Asian	2,804	4.40%	71.21	21.27	2,793	4.39%	69.46	20.07
	Hispanic	1,272	2.00%	63.79	21.27	1,285	2.02%	64.53	20.26
	Black	2,755	4.33%	56.51	21.94	2,747	4.31%	60.83	20.73
	White	55,653	87.40%	81.34	15.95	55,678	87.42%	79.94	16.55
		63,676	**100.00%**	**79.17**	**16.75**	**63,688**	**100.00%**	**78.07**	**17.07**
1999	Native American	1,157	1.78%	64.71	20.26	1,159	1.79%	68.41	19.71
	Asian	2,903	4.48%	73.04	19.88	2,904	4.48%	72.61	19.58
	Hispanic	1,337	2.06%	62.92	21.49	1,344	2.07%	66.85	21.54
	Black	3,148	4.86%	56.85	21.76	3,169	4.88%	63.28	22.01
	White	56,281	86.82%	81.38	16.19	56,302	86.78%	82.79	15.18
		64,826	**100.00%**	**79.14**	**16.89**	**64,878**	**100.00%**	**80.80**	**16.04**
2000	Native American	1,185	1.82%	66.32	19.00	1,210	1.86%	70.73	20.02
	Asian	3,069	4.73%	75.15	18.07	3,073	4.73%	76.4	17.87
	Hispanic	1,451	2.23%	65.28	19.29	1,442	2.22%	70.64	20.26
	Black	3,197	4.92%	60.48	19.51	3,215	4.94%	67.95	21.16
	White	56,045	86.29%	81.11	15.10	56,086	86.25%	84.46	14.36
		64,947	**100.00%**	**79.19**	**15.68**	**65,026**	**100.00%**	**82.70**	**15.22**
2001	Native American	1,186	1.86%	67.52	18.96	1,257	1.96%	72.06	20.45
	Asian	3,093	4.85%	75.60	17.28	3,110	4.86%	75.22	18.59
	Hispanic	1,590	2.49%	66.84	19.48	1,582	2.47%	69.55	21.74
	Black	3,618	5.67%	61.33	19.88	3,635	5.68%	66.58	22.28
	White	54,343	85.14%	82.24	14.72	54,439	85.03%	85.16	15.14
		63,830	**100.00%**	**80.08**	**15.41**	**64,023**	**100.00%**	**82.98**	**16.12**
Overalll	Native American	6,336	1.74%	66.40	19.31	6,312	1.75%	67.49	20.23
	Asian	16,191	4.44%	73.55	18.81	16,134	4.46%	71.16	19.96
	Hispanic	7,370	2.02%	65.25	19.86	7,306	2.02%	66.55	20.93
	Black	17,195	4.72%	58.92	20.18	17,153	4.74%	62.36	21.72
	White	317,381	87.08%	81.38	15.11	314,614	87.03%	80.66	16.28
		364,473	**100.00%**	**79.39**	**15.75**	**361,519**	**100.00%**	**78.85**	**16.94**

The overall results for the math test show a gap of 22.46 in percent correct scores between the highest group (Whites, 81.38%), and the lowest group (Blacks, 58.92%). Asians (73.55%) are second, followed by Native Americans (66.40%), and Hispanics (65.25%). Table 9 (p. 16) shows the effect sizes for performance differences between each pair of ethnic groups on the math test annually and overall. Overall, the Asian/White difference is half a standard deviation unit for the math test. For Native Americans versus Whites the effect size is 0.95. For Hispanics versus White students the difference is slightly higher (1.02). Finally, the difference between Black and White students is 1.43. To put this number in perspective, one should note that in the 1960s, the racial disparity in scores between Black and White students was approximately one standard deviation unit for a wide range of tests. After all of the efforts of the 70's and 80's to increase racial equity, this difference shrank to 0.75 of a standard deviation unit in the 1980s (Davenport, 1997). It is remarkable that the Black/White discrepancy in Minnesota for the past half-decade is almost twice the national 0.75 effect size seen in the 1980s, and is 50% higher than the one standard deviation unit obtained in the 1960s.

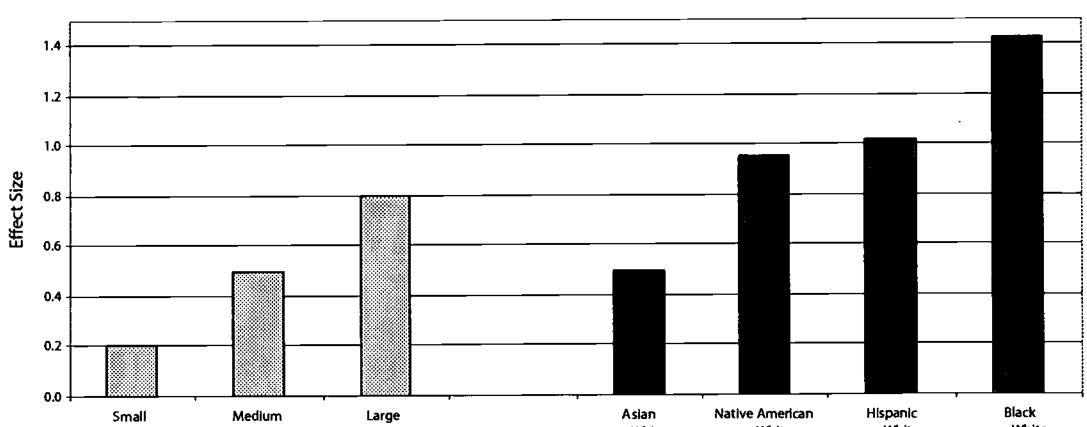
Table 8. Ethnic Groups' Representation in Minnesota: 1990 and 2000

	1990		2000	
	Counts	%	Counts	%
Native American	49,909	1.13	54,967	1.12
Asian	77,886	1.77	141,968	2.89
Hispanic	53,884	1.22	143,382	2.92
Black	94,944	2.15	171,731	3..50
White	4,130,395	93.72	4,400,282	89.58
Total	4,407,018	100.00	4,912,330	100.00

The category "Other" is not included in this table. 1990 data are from: http://factfinder.census.gov/servlet/BasicFactsTable?_lang=en&_vt_name=DEC_1990_STF1_DP1&_geo_id=04000US27; 2000 data are from: http://quickfacts.census.gov/qfd/states/27000.html

Figure 5 presents overall effect sizes for each ethnic group relative to White students. In addition to the results for ethnic groups, there are three additional effects classified as small/medium/large to aid in interpretation. The smallest of the ethnic effects is medium, for Asian students relative to Whites. For each of the other ethnic groups the effect size is at minimum large. A plot of the annual effect sizes for each ethnic group compared to Whites is given in Figure 6 (p. 17). Ideally, each of these lines would be near zero; and, if not, the next best condition would be

Figure 5
Overall Effect Sizes in Mathematics: Comparison of Minority and White Students



Note: the dark bars represent the size of the *difference* between the performance of minority groups and White students in mathematics. The lighter bars represent small, medium, and large effect sizes (differences) as described by Cohen (1988).

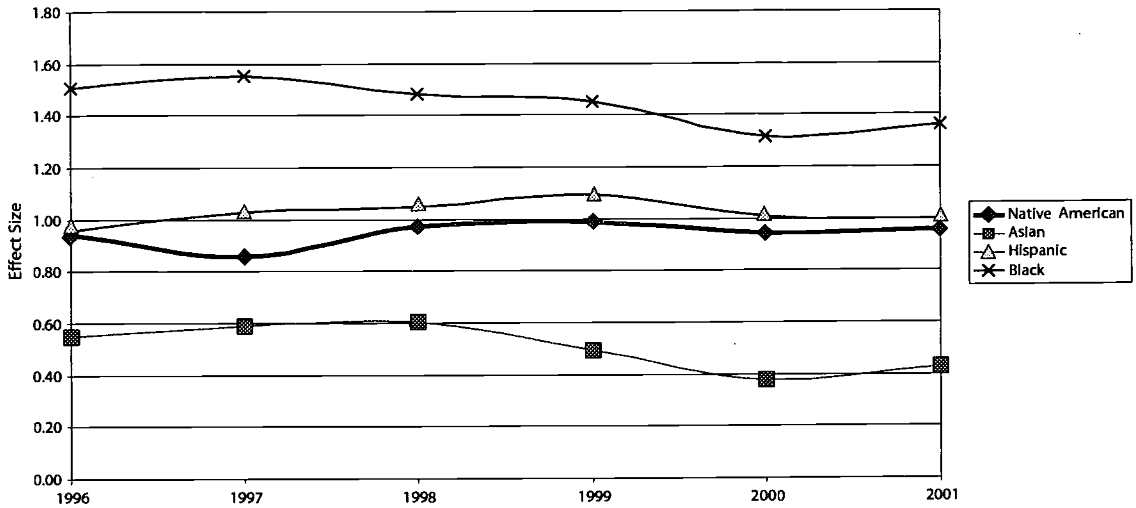
for the lines to show a downward trend, where the difference is moving toward zero. In Figure 6, the line for Blacks is at the top since they have the largest discrepancy compared to White students. While this line is far from zero, it exhibits the next best circumstance, a

Table 9. Effect Size Differences, by Ethnicity

		Mathematics				Reading			
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
1996	Native Americans	0.39	-0.02	-0.57	0.94	0.16	-0.08	-0.46	0.69
	Asian		-0.41	-0.96	0.55		-0.24	-0.62	0.53
	Hispanic			-0.55	0.96			-0.38	0.76
	Black				1.51				1.14
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
1997	Native Americans	0.27	-0.17	-0.70	0.86	0.07	-0.05	-0.43	0.77
	Asian		-0.44	-0.96	0.59		-0.12	-0.50	0.70
	Hispanic			-0.53	1.03			-0.38	0.82
	Black				1.55				1.20
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
1998	Native Americans	0.37	-0.08	-0.51	0.97	0.26	-0.03	-0.25	0.87
	Asian		-0.44	-0.88	0.60		-0.29	-0.51	0.61
	Hispanic			-0.44	1.05			-0.22	0.90
	Black				1.48				1.12
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
1999	Native Americans	0.5	-0.11	-0.47	0.99	0.26	-0.10	-0.32	0.90
	Asian		-0.60	-0.96	0.49		-0.36	-0.58	0.63
	Hispanic			-0.36	1.09			-0.22	0.99
	Black				1.45				1.22
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
2000	Native Americans	0.57	-0.07	-0.37	0.94	0.37	-0.01	-0.18	0.90
	Asian		-0.63	-0.94	0.38		-0.38	-0.56	0.53
	Hispanic			-0.31	1.01			-0.18	0.91
	Black				1.32				1.09
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
2001	Native Americans	0.52	-0.04	-0.40	0.96	0.20	-0.16	-0.34	0.81
	Asian		-0.56	-0.93	0.43		-0.35	-0.54	0.62
	Hispanic			-0.35	1.00			-0.18	0.97
	Black				1.36				1.15
		Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
Overall	Native Americans	0.45	-0.07	-0.47	0.95	0.22	-0.06	-0.30	0.78
	Asian		-0.53	-0.93	0.50		-0.27	-0.52	0.56
	Hispanic			-0.40	1.02			-0.25	0.83
	Black				1.43				1.08

Positive effect sizes occur when the ethnic group represented by the column has the higher mean. A negative value means that the ethnic group specified for the row obtained the higher score.

Figure 6
Year-to-year Trend of Effect Sizes for Mathematics: Minority Students Compared to Whites (1996–01)



downward trend. Since 1997, the difference has shrunk annually until 2001, when there was a slight upturn. However, the trend is still hopeful. The bottom line, showing data for Asian students, also exhibits a downward trend based on the last three years of testing. The 2001 results for Asian students exhibit a slight upturn, as does the line for Black students. The Hispanic and Native American lines show little movement in the desired direction. In fact, if the change in time were correlated with the change in effect size, the results would show that the greatest change in the desired direction is for Blacks: their correlation is –0.89. In other words, their effect size difference is decreasing over time. Asian students have a correlation of –0.38. This negative correlation indicates a general decrease in effect size over time. In contrast, the correlation for Hispanics (0.22) and Native Americans (0.43) are both positive, indicating a trend in the opposite direction: the effect sizes (or, the relative *differences* between Hispanic and Native American students compared to White students) are *increasing* over the years.

Table 1 (p. 7) shows the percentage of students in the lower scoring group **predicted** to score at or above the mean relative to students in the higher scoring group, assuming that the scores for both groups are normally distributed. In contrast, Table 10 (p. 18) shows the **actual** values based on the *BST* data, which are not normally distributed. Specifically, Table 10 shows the percentage of students in each ethnic group scoring at or above the mean for Whites. Note that approximately 60% of the White students score at or above their mean. Usually, one expects 50% of any group to score at or above its mean. However, this is true only when the scores follow a symmetric distribution. The *BST* scores in mathematics are negatively skewed. Most of the White students score fairly high; however, there are a few students with exceptionally low scores. These low scores lower the mean. The result is that more than half of the White students score above their mean.

Table 11 (p. 19) gives the actual distribution of scores relative to all students on the math test. (The score used here is number correct.) The score is based on the 68 items. Figure 7 (p. 18) shows the actual distribution of the mathematics scores for all students over the six years the (blue line) and the corresponding bell curve. Note that there are more lower scores than would be expected if the scores were to follow a standard bell curve. Also,

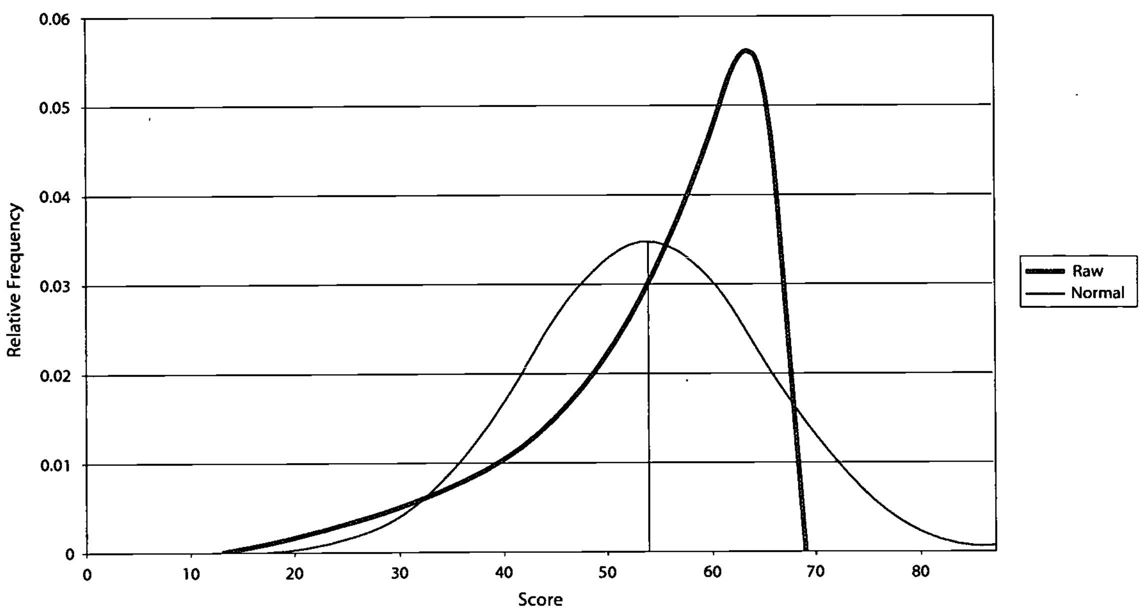
Table 10. Percentage of Students Scoring Over the Mean for Whites

		Native American	Asian	Hispanic	Black	White	Overall
1996	Mathematics	27.19	41.02	26.85	14.36	60.88	58.89
	Reading	31.63	37.44	28.99	19.31	56.64	53.38
1997	Mathematics	28.15	39.13	24.88	12.55	59.52	54.84
	Reading	29.52	34.45	29.33	18.80	58.13	53.87
1998	Mathematics	27.35	39.98	25.39	15.86	62.00	57.40
	Reading	28.78	39.74	29.88	23.41	63.58	59.27
1999	Mathematics	25.67	42.78	24.91	17.47	62.20	57.57
	Reading	25.28	36.57	27.53	22.40	60.45	56.08
2000	Mathematics	26.41	45.32	23.71	17.42	60.26	55.61
	Reading	31.65	43.28	32.80	28.15	65.68	61.06
2001	Mathematics	27.99	45.04	27.42	18.96	63.02	57.99
	Reading	30.31	34.89	28.70	23.77	61.64	56.66
Overall	Mathematics	26.53	42.07	25.25	16.22	60.76	56.19
	Reading	30.16	38.72	30.43	24.21	60.05	55.94

there are more scores above 56 than would be expected from the normal curve. Especially, note that the modal category—the one with the most subjects—is at 94.12% correct. This corresponds to 64 items correct out of 68—a very good score.

An in-depth look at Table 10 shows that, overall, 42.07% of Asian students scored at or above the mean for White students in mathematics. This percentage was 26.53 for Native Americans, 25.25 for Hispanics, and 16.22 for Blacks. The year-to-year results show that the percentage of Asian students scoring at or above the mean for Whites in mathematics ranged from 39.13% in 1997 to 45.32% in 2000. Their trend is generally upward. For Native Americans, the values ranged from 25.67% in 1999 to 28.15% in 1997. Native Americans seem to have had better relative performance early—in 1996 and 1997. Finally, their per-

Figure 7
Normal vs. Empirical Distribution for Mathematics



formance recovered slightly in 2001. The Hispanic group's performance varied, moving both upward and downward before ending in 2001 with their best performance relative to White students. The corresponding values for Black students ranged from a low of 12.55% in 1997 to 18.96% in 2001. In fact, the disparity in 1997 was so great that White students were almost 5 times as likely to have scores at or above their mean as Black students

Table 11. Score Distribution for the Mathematics Test: All Years

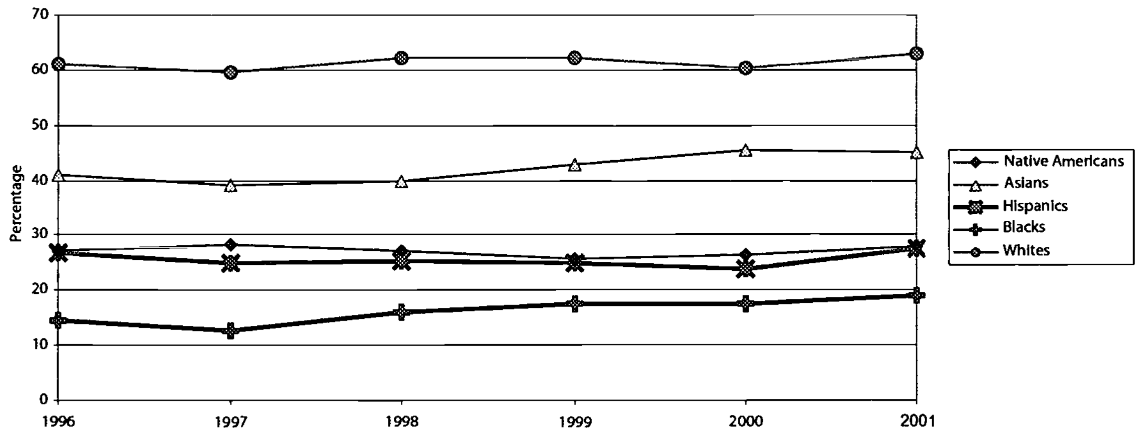
Score	%	N	% N	Score	%	N	% N
0	0.00	8	0.00%	35	51.47	2,797	0.76%
1	1.47	5	0.00%	36	52.94	2,990	0.81%
2	2.94	11	0.00%	37	54.41	3,251	0.88%
3	4.41	21	0.01%	38	55.88	3,454	0.93%
4	5.88	26	0.01%	39	57.35	3,757	1.01%
5	7.35	19	0.01%	40	58.82	3,790	1.02%
6	8.82	16	0.00%	41	60.29	4,112	1.11%
7	10.29	30	0.01%	42	61.76	4,516	1.22%
8	11.76	29	0.01%	43	63.24	4,741	1.28%
9	13.24	30	0.01%	44	64.71	5,228	1.41%
10	14.71	71	0.02%	45	66.18	5,646	1.52%
11	16.18	94	0.03%	46	67.65	6,059	1.64%
12	17.65	168	0.05%	47	69.12	6,616	1.79%
13	19.12	238	0.06%	48	70.59	7,201	1.94%
14	20.59	327	0.09%	49	72.06	7,650	2.07%
15	22.06	392	0.11%	50	73.53	8,406	2.27%
16	23.53	548	0.15%	51	75.00	8,724	2.36%
17	25.00	631	0.17%	52	76.47	9,461	2.55%
18	26.47	739	0.20%	53	77.94	10,315	2.79%
19	27.94	875	0.24%	54	79.41	11,092	3.00%
20	29.41	1,011	0.27%	55	80.88	12,152	3.28%
21	30.88	1,073	0.29%	56	82.35	12,917	3.49%
22	32.35	1,196	0.32%	57	83.82	14,016	3.78%
23	33.82	1,333	0.36%	58	85.29	14,944	4.04%
24	35.29	1,391	0.38%	59	86.76	16,338	4.41%
25	36.76	1,448	0.39%	60	88.24	17,595	4.75%
26	38.24	1,670	0.45%	61	89.71	18,739	5.06%
27	39.71	1,708	0.46%	62	91.18	19,871	5.37%
28	41.18	1,758	0.48%	63	92.65	20,557	5.55%
29	42.65	1,919	0.52%	64	94.12	20,669	5.58%
30	44.12	2,000	0.54%	65	95.59	19,333	5.22%
31	45.59	2,178	0.59%	66	97.06	16,506	4.46%
32	47.06	2,444	0.66%	67	98.53	11,430	3.09%
33	48.53	2,458	0.66%	68	100.00	4,921	1.33%
34	50.00	2,718	0.73%				

Score is the number of items answered correctly (0 - 68); % is the corresponding percent correct for a given number of items; N is the number of students with the given score; % N is the percentage of students with that score.

(59.52% versus 12.55%). By 2001, this number had decreased to slightly over three times as likely (3.32). The trend for Blacks is generally upward.

Figure 8 illustrates the results observed in Table 10. The top line is for White students, 60% of whom routinely score above their own mean. The next line represents Asians. The next two lines represent Native Americans and Hispanics, with the Native American score elevated slightly. The bottom line in the figure represents Black students. Note that while no

Figure 8
Year-to-year Trend of Minority Student Performance in Mathematics: Percentage Over the Mean for Whites (1996–01)

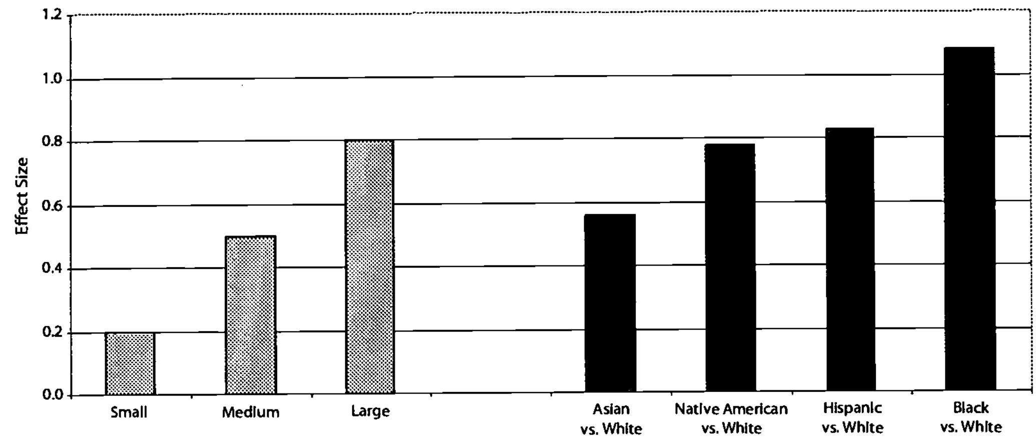


other group has fewer than 20% of the students in the group scoring at or above the mean for Whites in any of the six years, Black students never rise above this percentage.

Ethnic differences for reading are more moderate than those for math. Table 7 (p. 14) gives the annual and overall means for the five ethnic groups on the reading test. Table 9 (p. 16) shows the corresponding effect sizes. The overall difference for Asians versus Whites is 0.56 standard deviation units. This difference is medium and corresponds closely to the difference in math (0.50). Asians are the only group scoring relatively worse (compared to Whites) in reading than in math. The Native American versus White effect size for reading is 0.78, just under a large effect. This value is better than the corresponding value for math by 0.20. The Hispanic difference for reading is 0.83 (just over the threshold for a large effect)—a 0.2 decrease in effect size over their difference in mathematics. For Black students, the difference relative to White students is still large at 1.08. However, this value is a full 0.35 effect size points less than the difference in mathematics. Figure 9 (p. 21) shows ethnic differences for reading.

Although the reading differences are not as large as those for math (except for Asians), the annual trends for math suggest that more optimism is warranted about gap-closing in mathematics (two of four groups) than reading. The relative differences in reading performance do not systematically decrease for the ethnic groups. Asians had their best year compared to Whites in 1996, and matched that performance again in 2000. There was never a consistent 2-year trend in the six years of testing. In fact, the Asian group was 0.09 effect size points worse in 2001 (0.62) than in 1996 (0.53). Native Americans also had their best year in 1996 (0.69). Their relative performance then decreased over the next three years (0.77, 0.87, 0.90) before stabilizing for one year (0.90) and rising again in the final year (0.81). Native Americans ended 0.12 points worse in 2001 (0.81) than in 1996 (0.69).

Figure 9
Overall Effect Sizes in Reading: Comparison of Minority and White Students

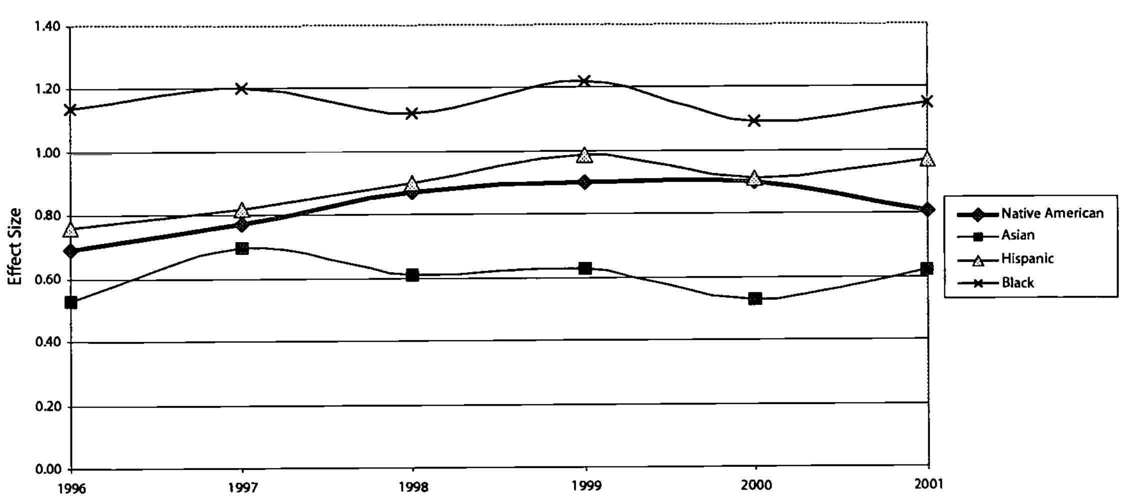


Note: the dark bars represent the size of the *difference* between the performance of minority groups and White students in reading. The lighter bars represent small, medium, and large effect sizes (differences) as described by Cohen (1988).

Hispanics also had their best performance in reading in 1996 relative to Whites (0.76). Their relative performance declined over the following three years (0.82, 0.90, 0.99) before improving in 2000 (0.91), only to dip again in 2001 (0.97). The difference between Hispanic students' scores and those of White students went from an effect size of 0.76 in 1996 to 0.97 in 2001. The best year for Blacks was 2000. As with the Asian group, there was no consistent 2-year trend for Blacks. Their ending point was 0.01 points worse than their start (1.14 versus 1.15). All groups had larger differences between themselves and the White sample in 2001 than in 1996, although this outcome could be the result of a biased selection of minority students being tested in 1996 and 1997 before the tests became mandatory.

Figure 10 shows effect size differences between specified ethnic groups and White students for reading. The higher the line, the greater is the distance between the performances. Blacks have the highest line, followed by Hispanics and Native Americans, whose lines are close. Asians (the lowest line) are closest to Whites. Reading results for the ethnic groups

Figure 10
Year-to-year Trend of Relative Effect Sizes for Reading: Minority Students Compared to Whites (1996–01)



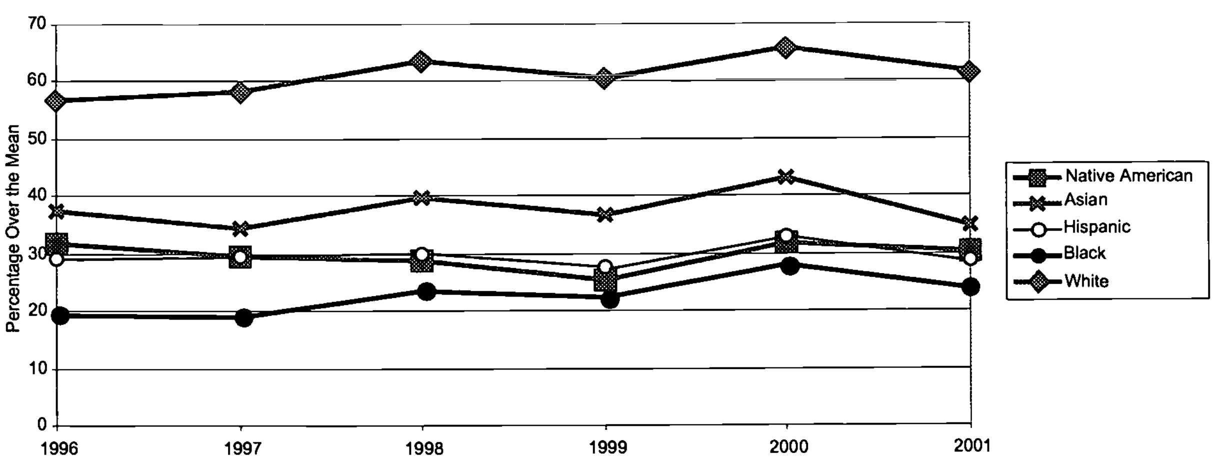
are in the same relative order as their mathematics results. Note that the ending point for each group is higher than the beginning point. This is because the performance gap is wider in 2001 than in 1996. Also note that the points for all of the groups go up and down: there is no consistent pattern of decrease in the ethnic reading gap for any ethnic group.

While it is difficult to identify optimistic trends in the reading data in Table 9 (p. 16) or Figure 10 (p. 21), two were found, using correlations of the effect size values over time. The results are the same as for math. The Asian and Black groups have inverse relationships between time and effect size, suggesting trends toward ethnic equity. The relationship relative to reading, however, is much weaker than for math, the correlation being only –0.03 for Asians and –0.20 for Blacks. Again, the Hispanic and Native American values are positive, suggesting increasing differences over time. This less favorable relationship is much stronger for reading than math. The correlations are 0.65 for Native Americans and 0.86 for Hispanics—suggesting that as time goes on, the differences (gaps) between minority ethnic groups and White students are increasing.

Looking again at Table 10 (p. 18), we see that overall, 38.72% of Asian students scored at or above the mean for White students in reading. The remaining percentages for reading are as follows: 30.16% for Native Americans, 30.43% for Hispanics, and 24.21% for Blacks. Relative to the annual results, Asian students ranged from 34.45% in 1997 to 43.28% in 2000. For Native Americans, values ranged from 25.28% in 1999 to 31.65% in 2000. For Hispanics, the value ranged from 27.53 in 1999 to 32.8 in 2000. The relative performance of Black students ranged from 18.8% in 1997 to 28.15% in 2000. The best relative performance for all groups appears to be in 2000, with a slight downturn for 2001.

Figure 11 shows the percentage of students in each ethnic category that scored at or above the mean for Whites for the reading test. The order is the same as for mathematics, with one exception. The White group is still first, followed by the Asian group. However, the Native American and Hispanic values are almost interchangeable. In fact the group with the higher value changes over the years, and changes more than once. Still, there is relative stability in the results—the points change little over the six years. The Black group is still

Figure 11
Year-to-year Trend of Minority Student Performance in Reading: Percentage Over the Mean for Whites (1996-01)

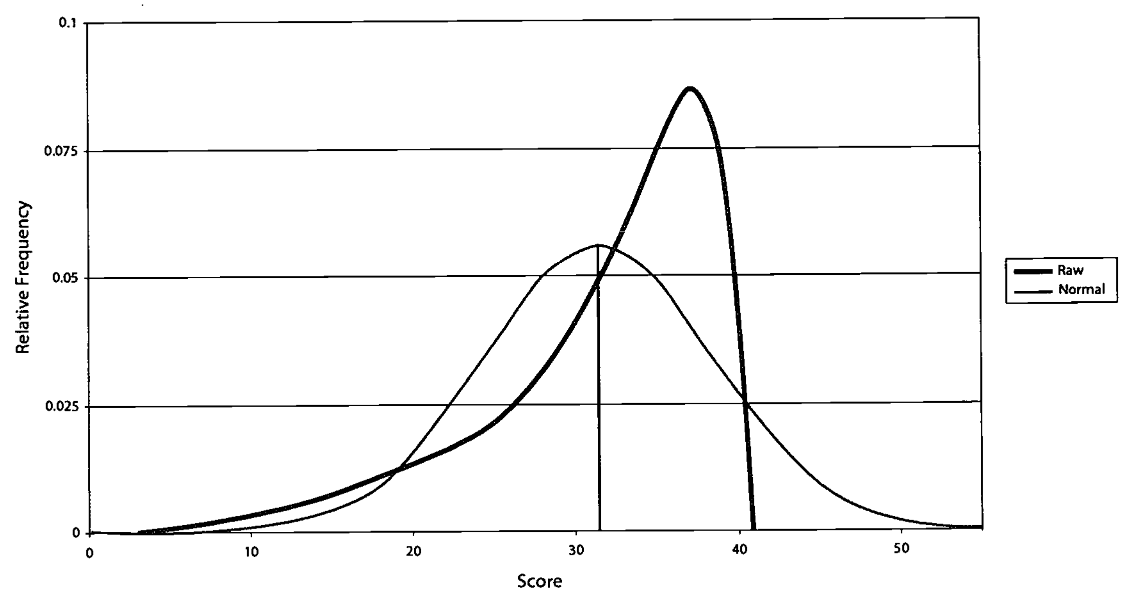


last, but the relative difference from the other groups is closer for reading than for mathematics.

Figure 12 shows results for reading that are very similar to the math results in Figure 7 (p. 18). As with math, the real distribution of reading scores is more spread out than would be expected if the scores followed a bell curve. Again, the scores are negatively skewed. Table 12 (p. 25) also gives the distribution of the reading scores. Again, the modal category is high: of a total of 40 items, 37 items correct is the most common score.

The summary statement for ethnicity is the same as for gender, but to a stronger degree: there are ethnic differences in performance on the *Basic Skills Tests*, and the discrepancy in results by ethnicity puts Minnesota at risk under President Bush's *No Child Left Behind* initiative, since closing gaps is a major component of that plan. While the White/Black and White/Asian gaps for math appear to be closing, the existing difference is large and progress is slow. In contrast to math, the differences in reading performance are smaller, but there are fewer trends toward ethnic equity.

Figure 12
Normal vs. Empirical Distribution for Reading



National Trends: A Comparison of Minnesota and U.S. Effect Sizes for Ethnicity
There is ample evidence that Minnesota is not alone in the presence and/or size of its ethnic differences. Lee (2002) uses SAT and NAEP data to show a decrease in ethnic differences during the 1970s and 1980s, but registers widening differences in the 1990s. Figure 13 (p. 24) reproduces information provided by NAEP (the National Center for Education Statistics [NCES], 2002) in its online data tool. This figure shows national score differences between Black and White 13 year olds for both mathematics and reading over the time period from 1969–99. Note that the score differences were dropping (moving toward equity) until 1989, when the differences began to increase annually, leaving equity behind.

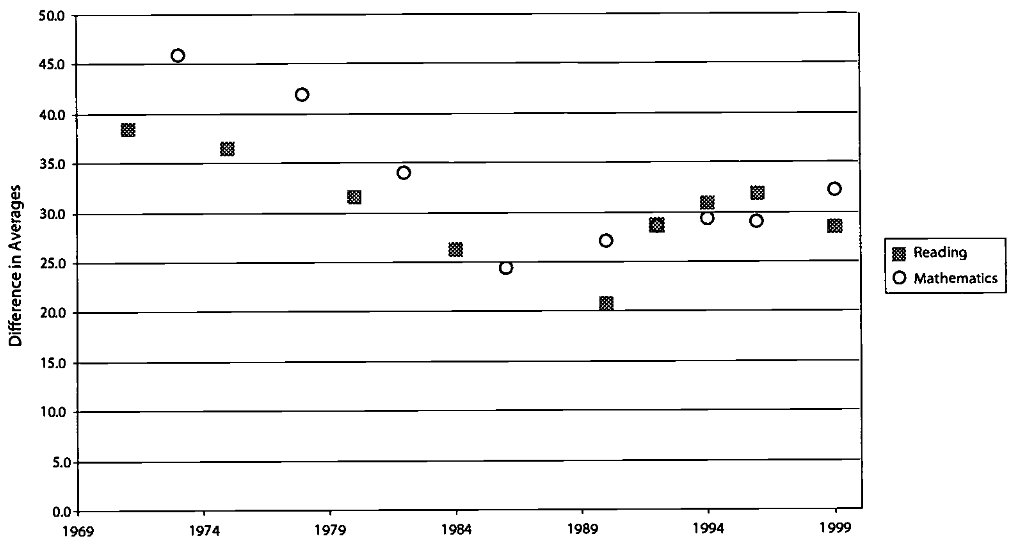
NCES (2000) also provides information so that one can calculate effect size values relative to the different ethnic groups for 13 year olds' reading and math tests over the past decade (this group is comparable to the 8th grade sample analyzed for the *BSTs*). These results are given in Table 13 (p. 26). Note that the Asian/Pacific Islanders' score each year is higher than the score for Whites in both mathematics and reading. The Asian/White difference is larger in mathematics and larger at the beginning of the 1990s. However, there is a trend toward ethnic equity in Asian and White students' performance on these NAEP assessments. The effect size difference in math is only 0.09 of a standard deviation unit in 2000, and the difference in reading for the last two tests (1994 and 1998) are miniscule (0.00 and 0.01).

Table 12. Score Distribution for the Reading Test: All Years

Score	% Correct	Count	% Students
0	0.0	14	0.00%
1	2.5	7	0.00%
2	5.0	23	0.01%
3	7.5	35	0.01%
4	10.0	70	0.02%
5	12.5	137	0.04%
6	15.0	245	0.07%
7	17.5	439	0.12%
8	20.0	712	0.19%
9	22.5	993	0.27%
10	25.0	1,299	0.35%
11	27.5	1,698	0.46%
12	30.0	2,119	0.58%
13	32.5	2,330	0.63%
14	35.0	2,680	0.73%
15	37.5	2,935	0.80%
16	40.0	3,245	0.88%
17	42.5	3,578	0.97%
18	45.0	3,906	1.06%
19	47.5	4,207	1.15%
20	50.0	4,561	1.24%
21	52.5	5,050	1.38%
22	55.0	5,519	1.50%
23	57.5	6,488	1.77%
24	60.0	7,168	1.95%
25	62.5	7,906	2.15%
26	65.0	9,091	2.48%
27	67.5	10,078	2.74%
28	70.0	11,456	3.12%
29	72.5	12,930	3.52%
30	75.0	14,903	4.06%
31	77.5	16,742	4.56%
32	80.0	19,226	5.24%
33	82.5	21,572	5.87%
34	85.0	24,563	6.69%
35	87.5	26,972	7.35%
36	90.0	29,594	8.06%
37	92.5	31,238	8.51%
38	95.0	30,848	8.40%
39	97.5	26,151	7.12%
40	100.0	14,460	3.94%

Similar to the *BST* results, NAEP also shows Blacks scoring the lowest of all of the defined ethnic groups in both mathematics and reading. Moreover, as with the *BST* data, the effect size difference for mathematics is larger than that for reading. Again, as with the state data, Blacks and Asians appear to be heading towards ethnic equity relative to White students. One difference here is that Asians are the highest scorers, and the relative gap with their White counterparts is closing. Another difference is that the gap for Blacks versus Whites is closing for both mathematics and reading, although, as with the *BST*, there is a stronger trend in mathematics. For Hispanics, the trends for both mathematics and reading are in the wrong direction—the gap between Hispanic and White students is widening over time. The data for Native Americans is not as clear. While the latest tests in both reading and mathematics show a widening of the performance gap between them and White students, this is just for one year. The preceding years show stability in the gap for reading (0.48 both years) and a decrease in the gap for mathematics (0.74 to 0.70 to 0.55).

Figure 13
Year-to-year Trend of Raw Score Differences
between White/Black Averages on the NAEP Assessments (1970–99)



The magnitude of the effect sizes for NAEP is smaller than for the *BST*, showing less ethnic disparity in the NAEP results. However, the NAEP results are far from desirable. While the Asian difference is miniscule for reading and goes from small to miniscule for math (Table 13, p. 26), the other ethnic groups show larger differences. Three of the effect sizes for Native Americans are medium; one is large and three are close to large. All seven of the effect sizes for Hispanics relative to Whites are large. Finally, all of the effect sizes for Blacks versus Whites in reading are large. The differences in mathematics are well beyond the threshold for a "large" effect, as defined by Cohen in 1988.

The bottom line is that the whole nation will be at risk of failing to meet adequate yearly progress goals under the provisions of the *No Child Left Behind* legislation. Olson (2002) states, "The fact that there's only one goal and every subgroup has to meet the same goal is probably the major problem." Predictions based on the current *BSTs* suggest that Minnesota will have difficulty closing ethnic gaps in mathematics, both because of their size and, in some cases, the direction of the trends. For the same reasons, the somewhat smaller ethnic gaps in reading will also be hard to close. Moreover, the national results offer little solace. While displaying smaller discrepancies, the trends were similar and the discrepancies, although smaller, are still large enough to defy easy solutions.

LEP Eligibility

Table 14 (p. 26) gives descriptive statistics for *BST* math and reading scores for students with limited English proficiency (LEP) versus those without LEP status. Between 1997 and 1998, when the *BST* became mandatory, there was almost a threefold increase in the number of students with LEP status taking the tests: from less than 600 students in 1997, the number rose to over 1600 in 1998. Also, the number of LEP-eligible students increased yearly after 1998. Table 15 (p. 27) shows the change in the percentage of LEP students tested from year to year. The percentage of LEP-eligible students tested actually dropped from 1996 to 1997 for both math (17.85% fewer) and reading (17.64% fewer). After the large increase in 1998, there is an additional 16.38% increase in LEP-eligible students taking the

Table 13. Minority Effect Size Differences for NAEP Data (Comparison with White Students)

		Native American	Asian/Pacific Islander	Hispanic	Black
Math	1990	0.74	-0.26	0.80	0.97
	1992	0.70	-0.33	0.94	1.24
	1996	0.55	*	0.94	1.19
	2000	0.90	-0.09	0.98	1.16
Reading	1992	0.48	-0.08	0.77	0.88
	1994	0.48	0.00	0.79	0.90
	1998	0.72	-0.01	0.84	0.87

A negative value means that the average score for Whites was smaller; There was no data for Asians/Pacific Islanders for the 1996 math test.

Table 14. Descriptive Statistics for Mathematics and Reading *Basic Skills Tests*, by LEP Eligibility

		Mathematics				Reading			
	LEP Status	N	%	Mean	SD	N	%	Mean	SD
1996	No LEP	56,153	98.76%	79.35	15.17	54,322	98.78%	72.32	17.73
	LEP	706	1.24%	57.41	17.11	669	1.22%	45.39	16.27
	Total	56,859	100.00%	79.08	15.19	54,991	100.00%	71.99	17.72
1997	No LEP	49,755	98.85%	79.95	15.34	48,362	98.87%	74.80	17.30
	LEP	580	1.15%	59.24	19.01	551	1.13%	48.05	17.65
	Total	50,335	100.00%	79.72	15.39	48,913	100.00%	74.50	17.30
1998	No LEP	62,052	97.45%	79.79	17.29	62,072	97.46%	78.70	17.36
	LEP	1,624	2.55%	55.35	21.44	1,616	2.54%	53.86	18.13
	Total	63,676	100.00%	79.17	17.41	63,688	100.00%	78.07	17.81
1999	No LEP	62,936	97.08%	79.80	17.52	62,992	97.09%	81.51	16.31
	LEP	1,890	2.92%	57.16	20.24	1,886	2.91%	57.17	19.08
	Total	64,826	100.00%	79.14	17.61	64,878	100.00%	80.80	16.90
2000	No LEP	62,857	96.78%	79.78	16.14	62,937	96.79%	83.39	15.34
	LEP	2,090	3.22%	61.24	18.95	2,089	3.21%	62.05	18.49
	Total	64,947	100.00%	79.19	16.23	65,026	100.00%	82.70	15.90
2001	No LEP	61,196	95.87%	80.79	15.92	61,396	95.90%	83.88	16.28
	LEP	2,634	4.13%	63.48	18.93	2,627	4.10%	61.97	19.67
	Total	63,830	100.00%	80.08	16.06	64,023	100.00%	82.98	16.43
Total	No LEP	354,949	97.39%	79.92	16.31	352,081	97.39%	79.42	17.22
	LEP	9,524	2.61%	59.64	19.76	9,438	2.61%	57.65	19.42
	Total	364,473	100.00%	79.39	16.41	361,519	100.00%	78.85	17.28

math test in 1999, and a corresponding increase of 16.71% for reading. From 1999 to 2000, the increase for both tests was almost 11%. The final year, 2001, shows a further increase of 26.03% for LEP students taking the math test and 25.75% for LEP students taking the reading test. The mean score for students with LEP status never reached the initial passing standard of 70% for students graduating in 2000. In fact, the highest mean percentage correct observed for LEP students on the mathematics test was 63.48% in 2001. The corresponding value was 62.05% for reading, and that occurred in 2000. The means for students classified as having limited English proficiency only reached 60% for math and reading in the final two years, 2000 and 2001. Table 16 gives the corresponding effect size data. As expected, LEP status seems to correspond more to reading scores. However, the difference between the overall effect size in mathematics and the effect size in reading is small (1.26 versus 1.24).

Table 15. Percentage Change in Number of LEP Test Takers from Year to Year (1996–01)

	Mathematics		Reading	
	Count	% Change	Count	% Change
1996	706	--	669	--
1997	580	-17.85%	551	-17.64%
1998	1,624	180.00%	1,616	193.28%
1999	1,890	16.38%	1,886	16.71%
2000	2,090	10.58%	2,089	10.76%
2001	2,634	26.03%	2,627	25.75%

Table 16. Effect Size Differences for LEP Eligible vs. LEP Non-eligible Students (1996–01)

	Mathematics	Reading
1996	-1.44	-1.52
1997	-1.35	-1.55
1998	-1.40	-1.39
1999	-1.29	-1.44
2000	-1.14	-1.34
2001	-1.08	-1.33
Total	-1.24	-1.26

Negative values represent a deficit in the LEP students.

Annually, with the exception of 1998, the relative reading performance of the LEP-classified students was worse than for math. In addition, the trend in both mathematics and reading suggests a consistent decrease in the performance difference between LEP-eligible and non-LEP-eligible students over the years. Math shows a drop from an effect size of 1.44 in 1996 to 1.08 in 2001. During this same period the difference in reading performance between LEP-eligible and non-LEP-eligible students decreased from 1.52 standard deviation units to 1.33. These results from 1996–01 show a proportionate decrease in the discrepancy for LEP-eligible versus non-LEP-eligible students in mathematics that is twice as large as for reading (1.44 – 1.08) / 1.44 = 0.25, versus (1.52 – 1.33) / 1.52 = 0.125.

Still, the performance differences between LEP-eligible and non-LEP-eligible students are large. They begin large—1.44 for math and 1.52 for reading, and end large—1.08 for math versus 1.33 for reading. However, while the differences are large, the trend is in the desired direction. It appears that limited English proficiency is becoming less of a barrier to student performance in both reading and mathematics, although it is becoming less of a barrier more rapidly in mathematics (see Figure 14).

Figure 14
Year-to-year Trend of Effect Sizes for Mathematics and Reading: LEP Students Compared to Non-LEP Students (1996–01)

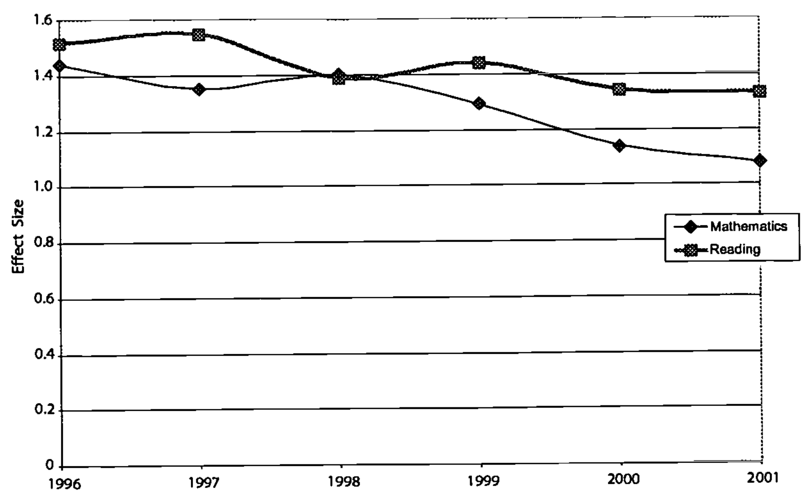


Figure 14 shows that, except for 1998, there is a consistent decrease in the gap between LEP-eligible and non-LEP-eligible students in both mathematics and reading. In 1998, there is an aberration in both scores: the math performance was relatively worse than expected and the reading performance was better than expected. Note too that the line for math (generally the bottom line) has a steeper slope and thus is changing (approaching equity) more quickly. The extra efforts expended on limited English proficient students appear fruitful, although the differences are still quite large, especially in reading.

While Minnesota is decreasing the gap between LEP-eligible and non-LEP-eligible students, there is still the question of time. If the change of 0.06 in effect size for mathematics between 2000 and 2001 were maintained, it would take 18 years to eradicate the remaining difference of 1.08 points. For reading, the task is even more daunting: given the rate of decrease (0.01 between 2000 and 2001), it would take another 133 years to eradicate the gap between the performance of students with and without limited English proficiency.

Table 17. Home Langage and LEP Status

	LEP vs Non LEP				Home Language: English vs. Other			
	LEP	Non-LEP	% LEP	Change	English	Non-Eng	% Non Eng	
1996	727	56,959	1.26%	--	55,659	2,027	3.51%	--
1997	587	51,115	1.14%	-0.099	49,356	2,346	4.54%	0.291
1998	1,800	63,130	2.77%	1.442	61,722	3,208	4.94%	0.089
1999	1,947	63,994	2.95%	0.065	62,553	3,398	5.15%	0.043
2000	2,286	64,656	3.42%	0.157	63,234	3,708	5.54%	0.075
2001	2,888	63,082	4.38%	0.282	61,748	4,222	6.40%	0.155

Given the prominence of LEP as one of the at-risk categories, we decided to extend our exploration of this variable by examining the relationship between LEP status, language most often spoken in the home, and performance on the math and reading portion of the *BSTs*. Table 17 shows the change in the state's population of LEP-eligible test takers and test takers for whom English is not the primary language at home (non-English homes). Note that with the exception of 1997, the number of students taking the *BSTs* with an LEP designator increases each year. Note too the ever increasing number of homes where English is not the primary language. While the number of students is rising in each category, the number of LEP test takers is increasing faster than the number of test takers for whom English is not the primary home language. The "Change" scores are the proportion of change in the percentage of students in the LEP category from one year to the next. For example, the –0.099 for LEP for 1997 suggests an almost 10% decrease in the proportion of LEP kids from 1996 to 1997 (1.3% to 1.1%). Again, with the exception of 1997, the change values for LEP are always larger than those for homes where English is not the primary language.

Table 18. LEP Status for Non-English Home Language

	No LEP	LEP	% LEP
1996	1,302	725	35.77%
1997	1,763	583	24.85%
1998	1,420	1,788	55.74%
1999	1,464	1,932	56.89%
2000	1,432	2,276	61.38%
2001	1,347	2,875	68.10%
Total	8,728	10,179	53.84%

These values suggest a relative increase in the number of students each year who are receiving LEP services. Table 18 (p. 28) contains only information about students whose primary home language is not English. The table shows that approximately one third of the test takers for whom English was not their primary home language received LEP services in 1996–97. By 2001, this ratio had risen to two thirds of those test takers—a substantial increase.

Given the above results for LEP and home language, one may wonder whether the increases in performance for the LEP-designated students result from real increases in their educational outcomes, or from merely classifying more students as having limited English proficiency. If the changes result from classification only, the average scores of the LEP-designated students would be likely to increase since the less proficient of the students whose home language is not English (if they were classified as having limited English proficiency) should not be as language-challenged as the majority of LEP-designated students. The mean of the LEP group would therefore be expected to increase. In addition, the average scores for non-LEP students whose primary home language is not English should also increase, since the latter group should be losing the more language-challenged students (and keeping the students who would be expected to perform better).

Table 19 shows summary statistics for three groups of students. The first

Table 19. Descriptive Statistics for Mathematics and Reading *Basic Skills Tests*, by Home Language and LEP Eligibility (1996–01)

		Mathematics				Reading			
		N	%	Mean	SD	N	%	Mean	SD
1996	Eng / No LEP	54,881	96.52%	79.52	15.10	53,108	96.58%	72.50	17.67
	No Eng / No LEP	1,272	2.24%	72.20	16.20	1,214	2.21%	64.38	18.49
	LEP	706	1.24%	57.41	17.11	669	1.22%	45.39	16.27
	Total	**56,859**	**100.00%**	**79.08**	**15.10**	**54,991**	**100.00%**	**71.99**	**17.67**
1997	Eng / No LEP	48,034	95.43%	80.31	15.13	46,648	95.37%	75.27	17.03
	No Eng / No LEP	1,721	3.42%	69.95	17.59	1,714	3.50%	62.02	19.49
	LEP	580	1.15%	59.24	19.01	551	1.13%	48.05	17.65
	Total	**50,335**	**100.00%**	**79.72**	**15.13**	**48,913**	**100.00%**	**74.50**	**17.03**
1998	Eng / No LEP	60,664	95.27%	79.96	17.21	60,681	95.28%	78.86	17.31
	No Eng / No LEP	1,388	2.18%	72.55	19.11	1,391	2.18%	71.65	17.99
	LEP	1,624	2.55%	55.35	21.44	1,616	2.54%	53.86	18.13
	Total	**63,676**	**100.00%**	**79.17**	**17.21**	**63,688**	**100.00%**	**78.07**	**17.31**
1999	Eng / No LEP	61,503	94.87%	79.90	17.48	61,560	94.89%	81.63	16.25
	No Eng / No LEP	1,433	2.21%	75.42	18.92	1,432	2.21%	76.13	18.04
	LEP	1,890	2.92%	57.16	20.24	1,886	2.91%	57.17	19.08
	Total	**64,826**	**100.00%**	**79.14**	**17.48**	**64,878**	**100.00%**	**80.80**	**16.25**
2000	Eng / No LEP	61,483	94.67%	79.85	16.10	61,564	94.68%	83.46	15.31
	No Eng / No LEP	1,374	2.12%	76.69	17.55	1,373	2.11%	79.92	16.16
	LEP	2,090	3.22%	61.24	18.95	2,089	3.21%	62.05	18.49
	Total	**64,947**	**100.00%**	**79.19**	**16.10**	**65,026**	**100.00%**	**82.70**	**15.31**
2001	Eng / No LEP	59,913	93.86%	80.87	15.88	60,110	93.89%	84.00	16.22
	No Eng / No LEP	1,283	2.01%	77.27	17.39	1,286	2.01%	78.47	17.93
	LEP	2,634	4.13%	63.48	18.93	2,627	4.10%	61.97	19.67
	Total	**63,830**	**100.00%**	**80.08**	**15.88**	**64,023**	**100.00%**	**82.98**	**16.22**
Total	Eng / No LEP	346,478	95.06%	80.07	16.24	343,671	95.06%	79.61	17.12
	No Eng / No LEP	8,471	2.32%	73.84	18.04	8,410	2.33%	71.79	19.38
	LEP	9,524	2.61%	59.64	19.76	9,438	2.61%	57.65	19.42
	Total	**364,473**	**100.00%**	**79.39**	**16.24**	**361,519**	**100.00%**	**78.85**	**17.12**

group's primary home language is English, and they do not have an LEP designation. The scond group's primary home language is not English, and they also are not classified as LEP. The final group of students is designated as LEP. The general trend for each of the three groups for both tests is positive—the scores are increasing. Note the similarity between both groups of students who are without an LEP designation. Note too that this similarity is increasing over time; the scores for the English in the home non-LEP group and the non-English in the home non-LEP group are getting closer. These results are consistent with the hypothesis that more mildly language-challenged students are being categorized as LEP as the years go by, pulling the weaker students out of the Non-English/Non-LEP group and leaving that group with more able students. Table 20 shows the corresponding effect sizes.

In Table 20 ,we see that the discrepancies in mathematics between the two groups without LEP designation (those whose primary home language is English versus those whose primary home language is not English) is decreasing with time, from an effect size of 0.43 in 1998 to 0.23 in 2001. Also note that for these two groups performance on the mathematics test is more similar than their performance on the reading test (with the exception of 1996 and 1998. The discrepancies in reading between these two groups also shrink until they undergo a substantial reversal in 2001. Of special interest is the fact that the LEP group's performance is also getting better relative to both of the other groups of subjects on both the mathematics and reading test. Again, we see that the relative performance for mathematics is better.

Table 20. Effect Size Differences, by LEP and Home Language

		Mathematics		Reading	
		No Eng/No LEP	LEP	No Eng/No LEP	LEP
1996	Eng/No LEP	-0.48	-1.46	-0.46	-1.53
	No Eng/NoLEP		-0.98		-1.07
1997	Eng/No LEP	-0.68	-1.39	-0.78	-1.60
	No Eng/No LEP		-0.71		-0.82
1998	Eng/No LEP	-0.43	-1.43	-0.42	-1.44
	No Eng/No LEP		-1.00		-1.03
1999	Eng/No LEP	-0.26	-1.30	-0.34	-1.51
	No Eng/No LEP		-1.04		-1.17
2000	Eng/No LEP	-0.20	-1.16	-0.23	-1.40
	No Eng/No LEP		-0.96		-1.17
2001	Eng/No LEP	-0.23	-1.10	-0.34	-1.36
	No Eng/No LEP		-0.87		-1.02
Total	Eng/No LEP	-0.38	-1.26	-0.46	-1.28
	No Eng/No LEP		-0.87		-0.83

Positive effect sizes occur when the ethnic group represented by the column has the higher mean; A negative value means that the ethnic group specified for the row obtained the higher score.

It is true that the general positive trends are consistent with the hypothesis that more students are being classified as LEP and thus, the No English/No LEP group benefits by ridding itself of the more language-challenged students. Similarly, the LEP group's increase in scores fits with this hypothesis, since the new group of students being classified are indeed performing at a lower level than the average No English/No LEP student, but better than the average LEP student. However, the continued increase in reading scores overall, and relative to these three groups, suggest that the improvement in reading performance for LEP-designated students is real. In contrast, the increase in scores for the three groups in mathematics suggests that the improved results are partially caused by a change in the way students are classified (since the *overall* math score is not rising). The results suggest that those who are classifying students are putting the most needy into the LEP-designated group. The additional services provided for English instruction appear to be having a positive effect, although improvements are coming slowly.

IEP Eligibility

Table 21 reports statistics for both the math and reading *BSTs* for students having individual education plans (IEP) versus those who do not. As with LEP status, the number of test takers with IEP status increased significantly between 1997 and 1998, when testing became mandatory. In 1998, the increase in the number of test takers with IEP status was slightly less than 40% for mathematics and slightly over 40% for reading. For 1999 and 2000, the increase in IEP students for both tests was approximately 5% per year. The number of IEP students stabilized in 2001; the numbers are not much different from those presented for 2000. Table 22 (p. 32) contains these results.

Table 21. Summary Statistics for IEP Students vs. Non-IEP Students

		Mathematics				Reading			
		N	%	Mean	SD	N	%		SD
1996	No IEP	51,272	90.17%	81.04	13.59	49,562	90.13%	74.21	16.33
	IEP	5,587	9.83%	61.09	18.82	5,429	9.87%	51.73	19.40
	Total	56,859	100.00%	79.08	14.19	54991	100.00%	71.99	16.66
1997	No IEP	45,133	89.67%	81.75	13.80	43,904	89.76%	76.72	15.85
	IEP	5,202	10.33%	62.07	18.42	5,009	10.24%	55.00	19.37
	Total	50,335	100.00%	79.72	14.35	48,913	100.00%	74.50	16.24
1998	No IEP	56,426	88.61%	82.00	15.05	56,455	88.64%	80.80	15.40
	IEP	7,250	11.39%	57.10	21.92	7,233	11.36%	56.78	20.77
	Total	63,676	100.00%	79.17	15.98	63,688	100.00%	78.07	16.10
1999	No IEP	57,225	88.27%	82.07	15.20	57,279	88.29%	83.52	14.05
	IEP	7,601	11.73%	57.09	21.84	7,599	11.71%	60.31	21.77
	Total	64,826	100.00%	79.14	16.12	64,878	100.00%	80.80	15.16
2000	No IEP	56,949	87.69%	81.97	13.98	57,001	87.66%	85.40	12.73
	IEP	7,998	12.31%	59.41	19.71	8,025	12.34%	63.50	21.83
	Total	64,947	100.00%	79.19	14.81	65,026	100.00%	82.70	14.17
2001	No IEP	55,853	87.50%	82.84	13.70	55,998	87.47%	85.95	13.71
	IEP	7,977	12.50%	60.70	20.35	8,025	12.53%	62.27	22.42
	Total	63,830	100.00%	80.08	14.70	64,023	100.00%	82.98	15.08
Total	No IEP	322,858	88.58%	81.97	14.27	320,199	88.57%	81.43	15.26
	IEP	41,615	11.42%	59.39	20.45	41,320	11.43%	58.92	21.54
	Total	364,473	100.00%	79.39	15.11	361,519	100.00%	78.85	16.10

For mathematics, the performance gap for students with an IEP classification versus those without begins in much the same way as the gaps related to LEP. However, whereas the LEP students are closing the gap with their non-LEP peers, the IEP students *increase* the discrepancy with their non-IEP peers before starting to close the gap. The gap is closing, but the end result for the IEP students is quite different from the LEP student results. Whereas the effect size between LEP-eligible versus non-LEP-eligible started within 0.03 points from the effect size of the IEP-eligible versus non-IEP-eligible for mathematics (1.44

for LEP versus 1.41 for IEP in 1996) [see Table 16, p. 27], their respective discrepancies ended 0.43 units apart (1.08 for LEP versus non-LEP and 1.51 for IEP versus non-IEP students in 2001).

Table 22. Percentage of Change in Number of IEP Students from Year to Year (1996–01)

	Mathematics		Reading	
	Count	% Change	Count	% Change
1996	5,587	–	5,429	–
1997	5,202	-6.89%	5,009	-7.74%
1998	7,250	39.37%	7,233	44.40%
1999	7,601	4.84%	7,599	5.06%
2000	7,998	5.22%	8,025	5.61%
2001	7,977	-0.26%	8,025	0.00%

For reading, the relative discrepancies were even more dissimilar. The effect size for LEP-eligible versus non-LEP-eligible students was 1.52 in 1996 (Table 16, p. 27). The corresponding effect size for the IEP-eligible and non-IEP-eligible was 1.35 (Table 23). Initially, the IEP-designated students are performing relatively better relative to the non-IEP designated students than the LEP-designated students compared to the non-LEP-designated students. While the LEP-eligible/non-LEP-eligible effect size decreased (1.52 to 1.33), the IEP-eligible/non-IEP-eligible effect size increased (1.35 to 1.57), leaving the IEP-designated students in a worse position than before.

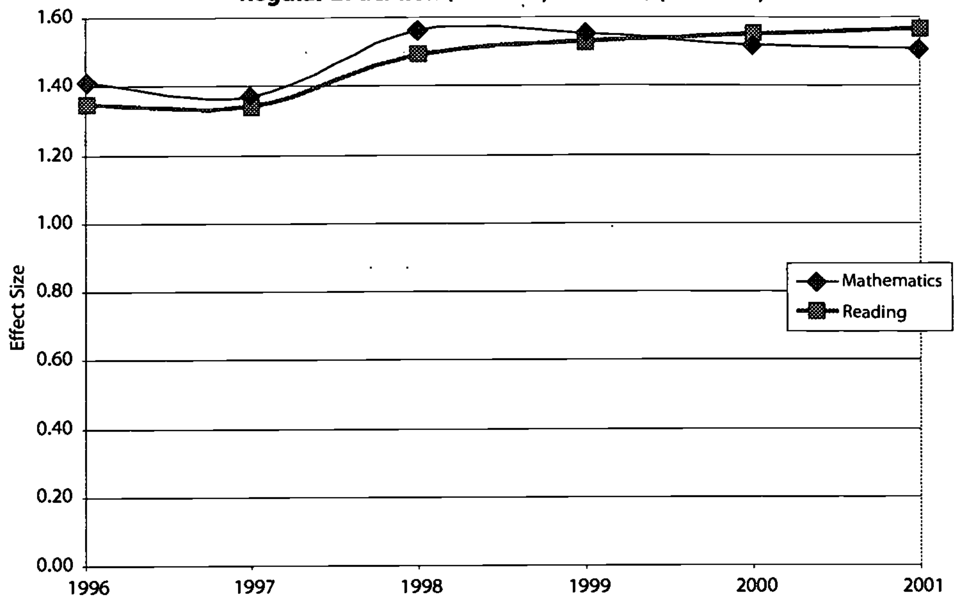
Not surprisingly, if we further parse the IEP/non-IEP results that were compared above, 8th graders with individual education plans have lower mean scores in both mathematics and reading than students without IEPs. Also, the mean scores for students with IEPs are, in both mathematics and reading, substantially lower than the passing standards. The mean percentage correct for the IEP sample ranged from the high 50's to the low 60's. Moreover, the average effect size, when comparing students with IEPs to students without IEPs, is rather large. The overall values given in Table 23 are 1.49 for mathematics, compared to 1.40 for reading. Although the differences are large, the large difference is not the major cause for alarm. What is alarming is that the trends, discernible from Table 23 and Figure 15 (p. 33), show that the relative performance of IEP versus non-IEP students is worse in 2001 than in 1996. The annual effect size values show an interesting pattern. The difference between students with IEPs and those without IEPs was larger for mathematics from 1996–99. For 2000 and 2001, the difference in reading is larger. However, if we assume that only the less severe IEP students were tested in 1996 and 1997, before testing became mandatory, the large increase in effect size from 1997 to 1998 in both math and reading might be explained by a change in the population of IEP students taking the tests. Starting with 1998, the trends show that the relative difference between IEP-eligible and non-IEP-eligible students is consistently decreasing for math while increasing for reading. This is true even though the mean for IEP students increases every year. The effect size also increases for reading because, even when the IEP students increase their scores, the non-IEP students increase their scores more. The final effect size values of 1.51 (math) and 1.57 (reading) are both large. However, over the years from 1998–01, the trend for mathematics is in the desired direction; by contrast, the trend for reading is not.

Table 23. Effect Size Differences for IEP Eligibility (1996–01)

	Mathematics	Reading
1996	-1.41	-1.35
1997	-1.37	-1.34
1998	-1.56	-1.49
1999	-1.55	-1.53
2000	-1.52	-1.55
2001	-1.51	-1.57
Overall	-1.49	-1.40

Negative values represent a deficit in the IEP students.

Figure 15
Year-to-year Trend of Effect Sizes for Mathematics and Reading: Special Education (IEP) Students Compared to Regular Education (Non-IEP) Students (1996–01)



IEP and LEP

The results for students who are both LEP- and IEP-eligible are even more pronounced than for students with either LEP or IEP status. Table 24 (p. 34) gives the overall and annual means for a) students with neither classification, b) students who are LEP-eligible but not IEP-eligible, c) students IEP-eligible but not LEP-eligible, and d) students both IEP- and LEP-eligible. Table 25 (p. 35) shows the corresponding effect sizes. While there were large differences between IEP-eligible and non-IEP- eligible students (Table 23, p. 32) and between LEP-eligible and non-LEP-eligible students (Table 16, p. 27), there is an even greater difference between students with neither classification and students who have both.

Table 25 (p. 35) shows the overall effect size differences for these two groups of students: 2.59 for mathematics and 2.39 for reading. Looking at the annual results, one finds that the smallest effect size difference in either mathematics or reading is the 1996 value for reading (2.23), and that subsequent differences oftentimes approach *three standard deviation units.* These differences are massive. Given an effect size of 2.5, as in Table 1 (p. 7) only 12% of these distributions would be expected to overlap. Furthermore, less than one percent of the students in the "Both" group would be expected to score at or above the mean for students not eligible for either classification. However, the actual data show that, overall, 3.2% of students with both IEP and LEP classifications (30 of the 945 students) scored at or above the mathematics mean for students with neither an IEP nor an LEP classification. The corresponding figure for reading was 3% (28 of 934 students).

Figure 16 (p. 36) shows the effect size differences between students with both IEP and LEP and students with neither, for mathematics and reading. The bars at the left side of the chart allow comparison of these effect size differences with small, medium, and large effects. The effect size differences for these two groups dwarf the bar showing a large

Table 24. Descriptive Statistics for Math and Reading *Basic Skills Tests*, by LEP & IEP Eligibility (1996–01)

		Mathematics				Reading			
	LEP/ IEP Status	N	%	Mean	SD	N	%	Mean	SD
1996	Neither	50,618	89.02%	81.34	13.29	48,946	89.01%	74.56	16.02
	LEP only	654	1.15%	58.18	17.20	616	1.12%	46.02	16.40
	IEP only	5,535	9.73%	61.21	18.83	5,376	9.78%	51.87	19.41
	Both	52	0.09%	47.74	12.44	53	0.10%	38.07	12.62
	Total	56,859	100.00%	79.08	13.97	54,991	100.00%	71.99	16.38
1997	Neither	44,598	88.60%	82.01	13.52	43,388	88.70%	77.06	15.53
	LEP only	535	1.06%	60.18	19.07	516	1.05%	48.84	17.54
	IEP only	5,157	10.25%	62.19	18.41	4,974	10.17%	55.13	19.34
	Both	45	0.09%	48.04	14.08	35	0.07%	36.43	15.20
	Total	50,335	100.00%	79.72	14.17	48,913	100.00%	74.50	15.98
1998	Neither	54,978	86.34%	82.65	14.31	55,015	86.38%	81.46	14.74
	LEP only	1,448	2.27%	57.52	21.01	1,440	2.26%	55.27	17.98
	IEP only	7,074	11.11%	57.59	21.83	7,057	11.08%	57.14	20.77
	Both	176	0.28%	37.58	16.00	176	0.28%	42.33	15.02
	Total	63,676	100.00%	79.17	15.50	63,688	100.00%	78.07	15.61
1999	Neither	55,535	85.67%	82.77	14.48	55,592	85.69%	84.26	13.19
	LEP only	1,690	2.61%	59.13	19.83	1,687	2.60%	58.92	18.70
	IEP only	7,401	11.42%	57.54	21.81	7,400	11.41%	60.80	21.71
	Both	200	0.31%	40.56	15.54	199	0.31%	42.39	15.61
	Total	64,826	100.00%	79.14	15.65	64,878	100.00%	80.80	14.58
2000	Neither	55,062	84.78%	82.62	13.33	55,113	84.76%	86.15	11.83
	LEP only	1,887	2.91%	62.95	18.48	1,888	2.90%	63.69	17.84
	IEP only	7,795	12.00%	59.78	19.67	7,824	12.03%	63.93	21.77
	Both	203	0.31%	45.34	15.66	201	0.31%	46.64	17.33
	Total	64,947	100.00%	79.19	14.42	65,026	100.00%	82.70	13.63
2001	Neither	53,488	83.80%	83.63	12.91	53,641	83.78%	86.93	12.53
	LEP only	2,365	3.71%	65.08	18.32	2,357	3.68%	63.56	19.19
	IEP only	7,708	12.08%	61.09	20.30	7,755	12.11%	62.77	22.39
	Both	269	0.42%	49.34	18.29	270	0.42%	48.08	18.32
	Total	63,830	100.00%	80.08	14.26	64,023	100.00%	82.98	14.40
Overall	Neither	314,279	86.23%	82.53	13.68	311,695	86.22%	82.04	14.67
	LEP only	8,579	2.35%	61.33	19.32	8,504	2.35%	59.10	19.14
	IEP only	40,670	11.16%	59.74	20.40	40,386	11.17%	59.26	21.52
	Both	945	0.25%	44.28	16.84	934	0.26%	44.47	16.84
	Total	369,888	100.00%	79.25	14.92	366,702	100.00%	78.70	15.89

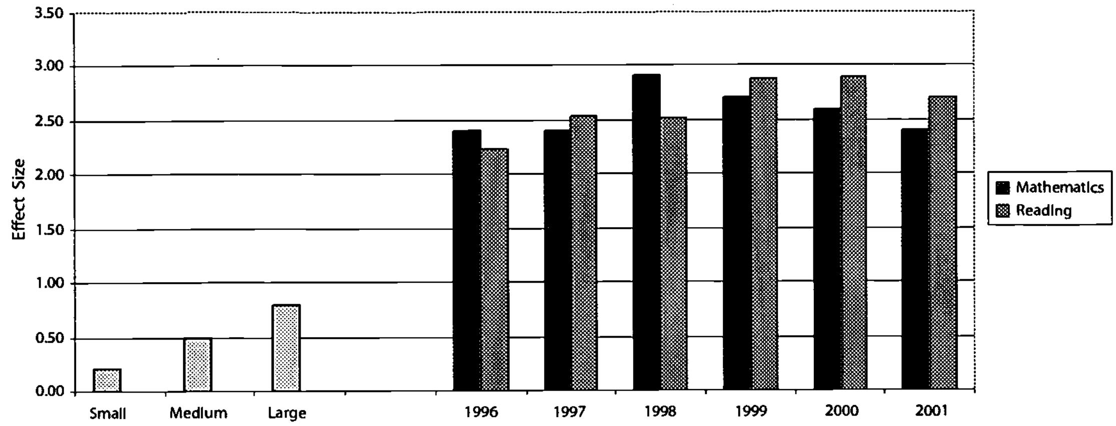
effect. In addition, the two smallest effect sizes for mathematics, and two of the three smallest annual effect sizes for reading for students with both IEP and LEP status, occurred during 1996 or 1997. Since the tests were not mandatory at that time, the data must be interpreted with caution: as with the data on students classified as having only LEP *or* IEP status, the effect sizes could be smaller merely because of differences in the population of test takers.

Table 25. Effect Size Differences for LEP & IEP Eligibility (1996–01)

		Mathematics			Reading		
		LEP only	IEP only	Both	LEP only	IEP only	Both
1996	Neither	-1.66	-1.44	-2.40	-1.74	-1.39	-2.23
	LEP only		0.22	-0.75		0.36	-0.49
	IEP only			-0.96			-0.84
		LEP only	IEP only	Both	LEP only	IEP only	Both
1997	Neither	-1.54	-1.40	-2.40	-1.77	-1.37	-2.54
	LEP only		0.14	-0.86		0.39	-0.78
	IEP only			-1.00			-1.17
		LEP only	IEP only	Both	LEP only	IEP only	Both
1998	Neither	-1.62	-1.62	-2.91	-1.68	-1.56	-2.51
	LEP only		0.00	-1.29		0.12	-0.83
	IEP only			-1.29			-0.95
		LEP only	IEP only	Both	LEP only	IEP only	Both
1999	Neither	-1.51	-1.61	-2.70	-1.74	-1.61	-2.87
	LEP only		-0.10	-1.19		0.13	-1.13
	IEP only			-1.08			-1.26
		LEP only	IEP only	Both	LEP only	IEP only	Both
2000	Neither	-1.36	-1.58	-2.59	-1.65	-1.63	-2.90
	LEP only		-0.22	-1.22		0.02	-1.25
	IEP only			-1.00			-1.27
		LEP only	IEP only	Both	LEP only	IEP only	Both
2001	Neither	-1.30	-1.58	-2.40	-1.62	-1.68	-2.70
	LEP only		-0.28	-1.10		-0.06	-1.08
	IEP only			-0.82			-1.02
		LEP only	IEP only	Both	LEP only	IEP only	Both
Overall	Neither	-1.44	-1.55	-2.59	-1.46	-1.45	-2.39
	LEP only		-0.11	-1.16		0.01	-0.93
	IEP only			-1.05			-0.94

The overall results in Table 25 show that students with only LEP status tend to score relatively better in mathematics than in reading (1.44 vs. 1.46 effect size) compared to students with neither LEP nor IEP status, although the difference is quite small. In contrast, the students with IEP status only score relatively better in reading than mathematics (1.45 vs. 1.55) compared to peers with neither LEP nor IEPs. Note, too, that the relative performance of students with IEP only, and students with both IEP and LEP differs in 1996 and 1997 from the remaining years. This finding tends to confirm that the population of IEP students who were tested in 1996 and 1997 was different from the population tested after the tests became mandatory. In contrast, the relative difference between students with LEP status only and students with neither classification appears to be consistent for all six years, suggesting more consistency in their performance over the entire six year period.

Figure 16
Year-to-year Trend of Effect Sizes for Mathematics and Reading: Comparison of Students with LEP and IEP vs. Students with Neither (1996–01)



Note: the bar heights represent the size of the *difference between* the performance of students with both IEP and LEP status, and those with neither in mathematics and reading. Bars representing small, medium, and large effect sizes (Cohen, 1988) are provided at the left of the chart, for comparison.

Although the number of LEP students changed drastically, their characteristics, as evidenced in the trends, appear consistent.

Specifically, the results from Table 25 (p. 35) show that for 1996 and 1997, students with IEP status only had better performance compared to the "Neither" group than did students with LEP only, both in mathematics (1.44 versus 1.66) and reading (1.39 versus 1.74). In 1998, however, the relative performances of LEP and IEP students compared to the Neither group were identical (1.62) for the LEP-only and IEP-only groups in mathematics. Thereafter, students with LEP only began closing their gap. The relative difference between IEP-only and 'Neither' groups has also been decreasing since 1998, but the rate of decrease is much slower. For reading, the LEP-only group started in 1998 with a larger difference compared to the Neither group (1.68) than the IEP-only group (1.56). However, the LEP-only group tended to close their performance gap with the Neither group, ending in 2001 with a modest decrease of 0.06, for an effect size difference of 1.62. On the other hand, the gap between the IEP-only group and the Neither group increased from an effect size of 1.56 in 1998 to 1.68 in 2001.

For students with both LEP and an IEP, educational equity is even less of a reality. Academically, these are Minnesota's neediest students. Their trend in mathematics since 1998 is in the right direction. The difference, however, between students with both IEP and LEP status and students with neither classification is huge. In reading, the difference is not only huge, but the trend is upward, with the gap increasing rather than decreasing.

Ethnicity Revisited

The interrelatedness between ethnicity and IEP/LEP status raises the question of what effect IEP/LEP status may have had on the ethnicity results presented earlier. We revisit the ethnicity results for students without IEP and LEP status in Tables 26 and 27 (p. 37). Table 26 shows the overall descriptive statistics reported by ethnicity for students without IEP or LEP status. Table 27 shows the corresponding effect size values. In Table 26, we see that the number of Native American students has decreased from 1.75% of the test takers (see Table 7, p. 14) to 1.59%. The number of Asian test takers has decreased from 4.45% to

3.17%. Hispanics were 2.02% of the general test taking population, where now they are at 1.43%. The percentages for Blacks go from a previous value of 4.73% to 4.10%. The only increase was in the percentage of White students. In the initial results (see Table 7, p. 14) they constituted approximately 87.05% of the students. Of students with neither IEP nor LEP status, Whites now comprise 89.70% of the sample. As expected, the mean values for

Table 26. Descriptive Statistics for Non-LEP and Non-IEP Students, by Ethnicity

	Mathematics				Reading			
	N	%	Mean	SD	N	%	Mean	SD
Native American	4,998	1.59%	70.47	17.38	4,963	1.59%	71.63	18.05
Asian	9,949	3.17%	80.47	14.96	9,910	3.18%	78.39	17.01
Hispanic	4,503	1.43%	72.58	17.00	4,497	1.44%	74.48	17.65
Black	12,864	4.09%	63.95	18.47	12,805	4.11%	67.65	19.69
White	281,965	89.72%	83.82	12.39	279,520	89.68%	83.13	13.69
Total	314,279	100.00%	82.53	12.95	311,695	100.00%	82.04	14.24

performance in Table 26 are higher than their corresponding values in Table 7. Also, the standard deviations for both tests for all ethnic groups are smaller here for this more homogenous group of students (Table 26) than for everyone (Table 7).

Table 26 also shows that, of this modified sample, the average score in both reading and mathematics for all ethnic groups except Blacks is higher than the passing mark of 70% set for students graduating in 2000. The overall mean for Asians and Whites is larger than the 75% standard required of students graduating in and after the year 2001. The overall mean in mathematics for Blacks is 63.95, and their mean score in reading is 67.65. For Asian students the effect size values relative to Whites are now better for both math and reading than previously (Table 27 vs. Table 9 [p. 16]). While the Asian/White mathematics effect size was 0.50 for all of the students, it is now 0.26 for the subset of students without LEP or IEP designation. The reading effect size was 0.56 and now is 0.33. The new effect sizes are

Table 27. Effect Sizes for Non-LEP and Non-IEP Students, by Ethnicity

	Mathematics				Reading			
	Asian	Hispanic	Black	White	Asian	Hispanic	Black	White
Native American	0.77	0.16	-0.50	1.03	0.47	0.20	-0.28	0.81
Asian		-0.61	-1.28	0.26		-0.27	-0.75	0.33
Hispanic			-0.67	0.87			-0.48	0.61
Black				1.53				1.09

also better for Hispanics versus Whites. Their math value was 1.02, but is now 0.87. Their reading value was 0.83 and is now 0.61. Looking only at students without IEP and LEP classifications, the performance of Asians and Hispanics becomes more comparable to that of White students.

This is not true for Black and Native American students. The overall effect sizes that we see when comparing the performance of Native American to White students is higher for mathematics than before (1.03 versus 0.95) and higher for reading (0.81 versus 0.78) [see Table 9, p. 16]. For Blacks, the new effect size for mathematics is higher (1.53 versus 1.43). The new value for reading is also higher (1.09 versus the previous 1.08). Thus, the modified samples of Native Americans and Blacks have relatively worse performance relative to Whites. The difference for mathematics, however, is more pronounced than the difference for reading.

Part of the reason for the effect size increases here for Blacks and Native Americans is that even though all ethnic groups showed improved performance after removing LEP and IEP

Table 28. Overall Raw Difference: Minority Student Scores Compared to White Student Scores

			Native American	Asian	Hispanic	Black	White
Math	Means	All Students	66.40	73.55	65.25	58.92	81.38
		Subset	70.47	80.47	72.58	63.95	83.82
		Difference	4.07	6.92	7.33	5.03	2.44
	White Difference	All Students	14.98	7.83	16.13	22.46	0.00
		Subset	13.35	3.35	11.24	19.87	0.00
Reading	Means	All Students	67.49	71.16	66.55	62.36	80.66
		Subset	71.63	78.39	74.48	67.65	83.13
		Difference	4.14	7.23	7.93	5.29	2.47
	White Difference	All Students	13.17	9.50	14.11	18.3	0.00
		Subset	11.50	4.74	8.65	15.48	0.00

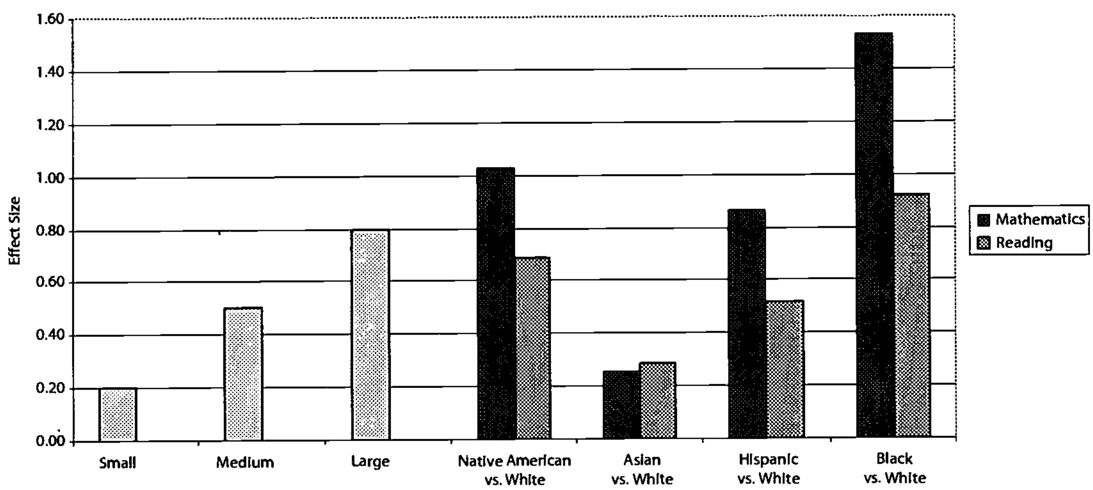
students from the sample, the populations also became more homogenous. Thus, the standard deviations were smaller. Since effect size is a function of the mean differences divided by the standard deviation, a drop in standard deviation can increase the effect size even if the mean difference stays the same or gets smaller. Table 28 parses this hypothesis.

The first two lines of data in Table 28 show the mean math scores for each of the ethnic groups. The first line is for all of the students (as in Table 7, p. 14) and the second is for a subset of students (as in Table 26, p. 38), excluding those having LEP or an IEP. The third line shows the increase in mean for the subset of students. The order of the gain (from low to high) is Whites, Native Americans, Blacks, Asians, and then Hispanics. Note that Native Americans and Blacks experience larger score gains than Whites. As expected, the means for the subset of students without LEP and/or IEP status are larger than for students overall. The next two lines of data show raw score differences in the mean score for each ethnic group compared to White students. In each instance the raw score difference for the subset of students is smaller than for all of the students. To illustrate, when we compare with all of the students, the difference in mean score for Blacks is 81.38 – 58.92 = 22.46. For the subset of students without LEP and IEP designation, the mean difference for Blacks is 83.82 – 63.95 = 19.87. Since the mean difference between the minority ethnic group and the White students is smaller for the subset than for all of the students, the increase in effect size for mathematics for Native American and Black students must be due to a smaller standard deviation for this subset of students.

Table 28 presents similar results for reading. Again, as expected, the mean reading scores for the subset of students without IEP and/or LEP designation is larger than for all of the students. The absolute difference in mean scores for each ethnic group compared to White students is smaller for the subset of students. The fact that the effect sizes are larger for two of the ethnic groups suggests that the standard deviations must be smaller. The bottom line is that the relative difference in mean scores is closer to that for Whites for each ethnic group; however, the subset of students is more homogenous, which leads to smaller standard deviations and larger effect sizes for two groups.

Figure 17 represents the effect size values for ethnic groups once IEP and LEP students have been removed. One of the first things to note in the figure is that the effect sizes in both math and reading for Asians compared to White students just exceeds the limit for a small effect. The Hispanic/White difference in reading is medium. The reading effect size for Native Americans versus Whites is between medium and large. The effect sizes for math for Hispanics and reading for Blacks just exceed the limit for a large effect. Finally,

Figure 17
Overall Effect Sizes in Mathematics and Reading:
Comparison of Minority and White Students Having Neither IEP or LEP Status



Note: the bar heights represent the size of the *difference between* the performance of minority students with neither IEP nor LEP status and White students, in mathematics and reading. Bars representing small, medium, and large effect sizes (Cohen, 1988) are provided at the left of the chart, for comparison.

there are only two effects larger than 1.0: for Native Americans in mathematics (1.03) and Blacks in mathematics (1.53).
In general, these results suggest less drastic ethnic differences than originally presented. However, the difference in mathematics for both Blacks and Native Americans compared to White students is now greater (in part, an artifact of the procedure used).

Free or Reduced Price Lunch Eligibility
Our final set of data compares students who are eligible for free or reduced-price lunch (a proxy for low socio-economic status [SES]) with students who are not eligible. Results are

Table 29. Descriptive Statistics for Mathematics and Reading *Basic Skills Tests*, by Free/Reduced-price Lunch Eligibility (1999–01)

		Mathematics			Reading		
		N	Mean	SD	N	Mean	SD
1999	No F/R Lunch	48,627	82.82	15.17	48.636	84.18	14.04
	F/R Lunch	16,199	68.11	21.10	16,242	70.69	20.35
	Total	64,826	79.14	16.85	64,878	80.80	15.86
2000	No F/R Lunch	49,459	82.30	14.26	49,520	85.62	13.31
	F/R Lunch	15,488	69.24	19.25	15,506	73.37	19.49
	Total	64,947	79.19	15.60	65,026	82.70	15.90
2001	No F/R Lunch	48,200	83.42	13.85	48,302	86.37	14.01
	F/R Lunch	15,630	69.76	19.20	15,721	72.57	20.72
	Total	63,830	80.08	15.34	64,023	82.98	15.92
Overall	No F/R Lunch	146,286	82.84	14.45	146,458	85.39	13.82
	F/R Lunch	47,317	69.02	19.90	47,469	72.19	20.23
	Total	193,603	79.46	15.95	193,927	82.16	15.63

presented in Tables 29 and 30. (This information was unavailable for 1996 and 1997, and results for 1998 were omitted.) Eighth graders who are eligible for free or reduced-price lunch have lower mean scores and their scores vary more (i.e., they have higher standard deviations) in both math and reading, compared to students not eligible for free or reduced-price lunch. Students eligible for free or reduced-price lunch had scores around 69% on the mathematics test, just a little short of the 70% standard set for the graduating class of 2000. Their mean reading score was somewhat higher, around 72%. While the overall trend is for reading scores to increase, this is not true for students who are eligible for free or reduced-price lunch. Their mean reading score goes up from 1999 to 2000, and down again from 2000 to 2001. On the other hand, the mean math score for these students consistently increases from 1999 to 2001.

Table 30. Effect Size Differences for Free/Reduced-price Lunch Eligibility (1999–01)

	Math	Reading
1999	-0.87	-0.85
2000	-0.84	-0.77
2001	-0.89	-0.87
Overall	-0.87	-0.84

Negative values indicate that the Group on Free/Reduced Price Lunch scored lower.

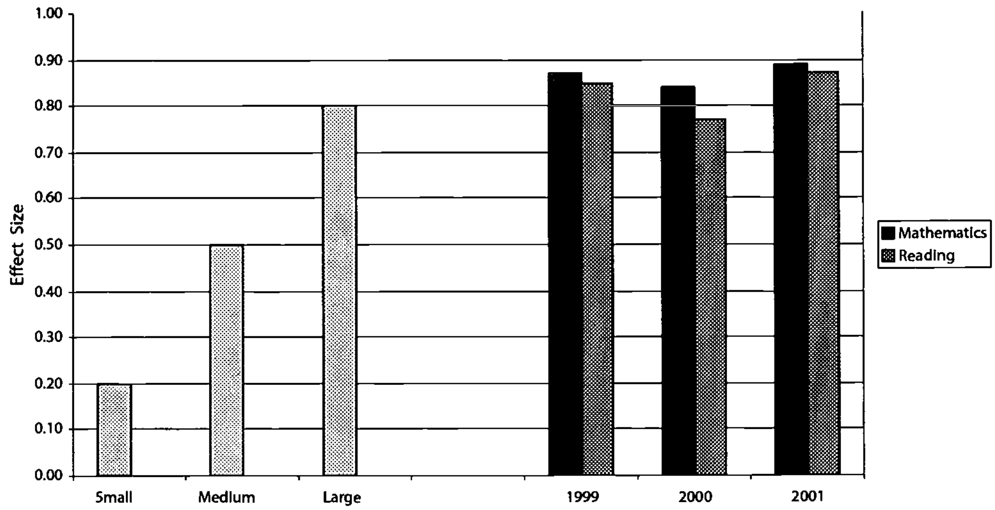
Overall, the effect size difference in math, 0.87, is slightly larger than that for reading, 0.84. Both of these values fall just over the 0.80 threshold for a large effect. The effect sizes in math are consistently larger than those for reading across all three years, and neither the trend for math nor the one for reading is positive. The largest difference between low and high socio-economic status (SES) students occurs in 2001 on both tests.

The difference between these groups appears to be widening. Figure 18 (p. 41) illustrates the effect size differences. Each of the effects is close to the standard large effect threshold (0.80). Only one of the

Table 31. Effect Sizes for NAEP Data: Free/Reduced-price Lunch- Eligible vs. Ineligible

	Year	Effect Size
Math	1996	0.82
	2000	0.87
Reading	1998	0.72

Figure 18
Effect Sizes for Mathematics and Reading: Comparison of Students Eligible for Free/Reduced-price Lunch and Students Not Eligible (1999–01)



Note: the bar heights represent the size of the *difference between* the performance of students eligible for free or reduced-price lunch vs. those not eligible, in mathematics and reading. Bars representing small, medium, and large effect sizes (Cohen, 1988) are provided at the left of the chart, for comparison.

effects is smaller than 0.80 (reading, in 2000). Finally, the bars for the 2001 results are the tallest, indicating larger discrepancies.

National Trends: A Comparison of Minnesota and U.S. Effect Sizes for Socio-economic Status

NCES (2002) also has national data for 13 year olds on reading and mathematics, categorized by whether the student is eligible for free or reduced-price lunch. Table 31 (p. 41) contains those results. As for the *BSTs*, this data also shows math performance to be relatively worse. The magnitude of the national effects is very similar to that for the *BST*; however, the national values are just a bit smaller. Thus, Minnesota and the nation face a similar problem in raising the scores of low-income students. The national trend for math suggests a widening of the performance gap for mathematics. There was only one year of data available for reading and thus, no trend could be discerned.

CONCLUSIONS

Given that the Minnesota *Basic Skills Tests* have been in place for over half a decade, it is reasonable to desire information on emerging trends. This manuscript began with an explanation of the philosophical foundations for Minnesota's graduation rule. It then provided a historical look at the math and reading portion of the Minnesota *Basic Skills Test* and identified trends relating to time, gender, ethnicity, limited English proficiency status, individual educational plan status, and socio-economic status. Data were obtained on the reading and mathematics portions of the Minnesota *Basic Skills Test* for 8th grade students taking the regular administration of the test during the years 1996–01. The basic statistic used was the effect size, since it gives the distance between mean scores for two groups in terms of their standard deviation, so that each comparison is on a standard scale. Finally, the results are put into the context of national achievement results (NAEP) and President Bush's *No Child Left Behind* educational initiative. The main results are as follows.

Data for the first two years of testing, 1996 and 1997, may not be comparable to data for the final four years (1998–01) due to a change in the population of students taking the exam. The Minnesota *Basic Skills Test* was not mandatory the first two years, and students could find alternative means to satisfy this testing requirement. After 1997, the *BST* was the only test that could be used to satisfy the requirement.

Students' performance in reading appears to be improving over time. This increase, however, appears to be leveling off. Students' performance in mathematics has no clear pattern over the period studied. However, the difference between the last two years hints at improvement.

Gender differences in mathematics favor males. However, this difference has consistently decreased over time for the *BSTs* and has almost vanished. This finding is consistent with the goal of closing achievement gaps as outlined in President Bush's *No Child Left Behind* act. Gender differences on the *BST* in reading are much larger than the gender differences in mathematics, and these differences favor females. In contrast to the math differences, differences in reading performance are not decreasing over time. The difference in 2001 is the same as the difference in 1996. NAEP results also show that differences for reading favor females and differences in mathematics favor males. The absolute differences were larger for NAEP, indicating a larger gender difference nationally than for Minnesota, in both math and reading.

Consistent with census findings, the relative percentage of White students in the student population is decreasing over time, as Minnesota becomes a more diverse state. The percentage of students in each minority group is rising except for Native Americans (this percentage remained virtually the same from 1990 to 2000).

Performance among ethnic groups is in the following order (from high to low): Whites, Asians, Native Americans, Hispanics, and Blacks. The performance of the Asian group places them approximately halfway between the White group and the Native American group. Moreover, the performance for the Native American and Hispanic groups are relatively close. Finally, the Black group is by itself at the bottom. The difference in performance between Black and White students is huge, especially in mathematics. While there is a trend towards ethnic equity in math for Blacks and Asians, the gap for Black students is enormous. A much weaker trend is present for Hispanic and Native American students, for whom the gaps are widening, not closing.

In reading, there is a strong trend away from ethnic equity for Native Americans and Hispanics. A much weaker trend for Blacks and Asians is in the desired direction. Based on the ethnic differences in scores on the *BSTs*, ethnic gaps in performance on mathematics tests will be hard to close: first, because of their magnitude and second, because of the strength and direction of the trends. The gaps in reading performance that relate to ethnicity will also be hard to close, due first to the strength and direction of the trends and then due to their size.

From a historical perspective, NAEP data shows a decrease in performance differences between Black and White students from 1970 until approximately 1989, when the gap began widening again. Effect size results from NAEP for 13 year olds over the last decade show a widening of the racial gap between Whites and both Native Americans and Hispanics. The gap between Blacks and Whites is closing, but it is the largest of the gaps.

Finally, the Asian/White gap is closing, but the national Asian group started with scores *higher* than for Whites. Ethnic differences appear to be real: sometimes they are large, and at times the trend is in the wrong direction. As opposed to gender, the differences among ethnic groups as manifested by scores for the *BST*s are larger than those shown in the national results.

Students with limited English proficiency do less well in both reading and mathematics than students without this classification. As expected, these students score relatively worse in reading. However, the difference between deficits for reading and mathematics is small. Minnesota is decreasing the performance gap between LEP-eligible and non-LEP-eligible students in both mathematics and reading. The LEP system appears to be working; however, the size of the disparity and the rate of decrease will make this a long-term endeavor. Note, however, that part of the positive results for mathematics may be due to a change in classification of LEP students.

Students with individual education plans score less well in both reading and mathematics than students without this classification. The discrepancies are almost twice as much as Cohen (1988) defined for a large effect. In general, the discrepancies for IEP-eligible versus non-IEP-eligible students are larger than those for LEP-eligible versus non-LEP-eligible students. Over the last four years of testing (1998–01), the trend for IEP-eligible versus non-IEP-eligible students in mathematics is in the correct direction, although the ending difference (2001) is still huge. The trend for reading, however, shows ever-widening gaps.

Minnesota's academically neediest students, those with both limited English proficiency and individual educational plans, are performing at extremely low levels. The effect size difference between students classified with both LEP and IEP versus students without either of these classifications is close to 2.5 standard deviation units for both mathematics and reading. The trend for equity between students with both IEP and LEP versus those without either classification is in the right direction for mathematics. However, the gap is still huge. The 2001 effect size is 2.4. For reading the trend is not favorable. There is a consistent rise in the effect size difference (away from equity), reaching almost 3 standard deviation units in 2000 before falling to 2.7 in 2001. However, the sample size for the sudents designated as both IEP and LEP is small, limiting generalization of the results.

Because of the interrelatedness of ethnicity, IEP, and LEP, another analysis was conducted to compare ethnic differences for students who had neither IEP nor LEP classification. The results show a decrease in the effect size difference between White versus Asian and Hispanic students. In contrast, the effect sizes for Blacks and Native Americans relative to Whites were larger. The increase in effect size for these latter two groups, however, was partially an artifact of the comparison procedure.

Students eligible for free or reduced-price lunch score worse than their counterparts who are not eligible. The difference is close to 0.80—a large effect. Although the discrepancies are close, they are a little larger for mathematics. The Minnesota data suggest that the SES gap appears to be widening for performance in both mathematics and reading. The trend and size of the effect are consistent with those seen in the NAEP data. There, the effect sizes are also close to 0.8 and for mathematics, the NAEP gap is also widening.

Minnesota and the nation have a large task ahead of them—a task that has, up to now, never been successful. That task is the closing of achievement gaps for various subgroups of the population. There are large and real gender differences, especially in reading. The differences between White students and Native Americans, Hispanics, and Blacks is much too large. The difference for Blacks is even larger than in the 1980s—a conclusion confirmed by multiple data sources. While the gaps are decreasing for LEP students compared to non-LEP students, the differences are still huge, especially for reading. It is not likely that this gap will be eradicated in the near future. For IEP students, as compared to non-IEP students, the gap is even larger than those between LEP and non-LEP students, and the IEP trends are in the wrong direction (the gap is increasing). Finally, socio-economic status is also related to performance differences. Five out of six of the effect sizes relating to SES were at least large, and the trend is in the direction of increasing difference (Hart, 1995).

Of these trends, the most informative may be the slow closing of the gap between LEP-eligible and LEP non-eligible students, especially in reading and partially in mathematics. This closing of the gap is clearly associated with participation in a particular type of educational intervention, English as a Second Language (ESL) programs. This result points to the importance of English language proficiency for the development of both reading and mathematics skill, particularly when the mathematics skill involves applied word problems. If improved English language proficiency is essential to the performance of students from homes in which the primary language is other than English, it may also be important for students whose home and community environments do not provide a rich background in the kind of "standard" or "academic" English used in school textbooks and reading materials. According to some, oral language development provides the foundation for later school success, particularly in reading, and homes vary substantially in the language experiences they provide to children.

Minnesota and the nation have much work to do to close achievement gaps. In addition to being "the right thing to do," the United States cannot afford to educate only a portion of its citizens. Our world is now in the information age and knowledge is capital. The U.S. needs an educated citizenry in order to be competitive in this global marketplace. If changes are not made, *No Child Left Behind* will become just another slogan: the genesis of a new wave of failed public policies aimed at educating our nation's "at-risk" students.

REFERENCES

Bush, G. W. (2001). *No child left behind.* Retrieved in July 2002 from: http://www.whitehouse.gov/news/reports/no-child-left-behind.pdf

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Davenport, Jr., E. C. (1997) Testing and bias. *Colors, 6,* #3. Minneapolis, MN: Four Colors Productions, Inc.

Davison, M.L., Erickson, R.N., Davenport, E., Kwak, N., Irish, M.L., Bielinski, J., Danielson, H., Kim, S-K., Seo, Y.S., Smith, M.J., Wick, S. (1999). *1998 Minnesota education yearbook: The status of pre-k-12 education in Minnesota.* Minneapolis, MN: Office of Educational Accountability, College of Education and Human Development, University of Minnesota.

Davison, M. L., Schleisman, J. L., Koeppen, L., Wu, Y-C, & Kwak, N. (2001). *A few weeks in summer: Post-summer school achievement among state-funded students who do not initially pass Minnesota's high school graduation test.* Office of Educational Accountability, College of Education. University of Minnesota.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in everyday experiences of young American children.* Baltimore, MD: P. H. Brookes.

House of Representatives (2001). *No child left behind act of 2001.* 107th Congress 1st Session, conference report 107-334 to accompany H.R. 1.

Lee, J. (2002). Racial and ethnic achievement gap trends: Reversing the progress toward equity? *Educational Researcher, 31,* 3-12.

National Center for Education Statistics (2002) NAEP data tool. Available online at: http://nces.ed.gov/naep3/naepdata/search.asp.

Olson, L. (2002). 'Inadequate' yearly gains are predicted. *Education Week, 21,* April 3, 2002.

Roy Wilkins Center for Human Relations and Social Justice (2000, October). *Racial disparities in Minnesota Basic Skills Test scores, 1996-2000.* Minneapolis, MN: University of Minnesota.

Roy Wilkins Center for Human Relations and Social Justice (1997, March). *Summary of results: An analysis of the 1996 Minnesota basic skills test scores.* Minneapolis, MN: University of Minnesota.

Schleisman, J. L., Peterson, K. A., and Davison, M. L. (2000). *Back to the basics: An investigation of school-and district-level remediation efforts associated with Minnesota's basic standards for high school graduation.* Minneapolis, MN: Office of Educational Accountability, College of Education and Human Development, University of Minnesota.

U.S. Census Bureau. American Fact Finder table DP-1. General Population and Housing Characteristics: 1990, Data Set: 1990 Summary Tape File 1 (STF 1) - 100-Percent data, Geographic Area: Minnesota. Retrieved in July 2002 from: http://factfinder.census.gov/servlet/BasicFactsTable?_lang=en&_vt_name=DEC_1990_STF1_DP1&_geo_id=04000US27.

U.S. Census Bureau. State and County QuickFacts for Minnesota: People QuickFacts. Retrieved in July 2002 from: http://quickfacts.census.gov/qfd/states/27000.html.

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Urban Education
Box 40, Teachers College, Columbia University
New York, NY 10027

Telephone: 212-678-3433
Toll Free: 800-601-4868
Fax: 212-678-4012

WWW: http://eric-web.tc.columbia.edu

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org