Analysis of the 1996 Minnesota Basic Standards Test Data

Roy Wilkins Center for Human Relations and Social Justice
Hubert H. Humphrey Institute of Public Affairs
University of Minnesota

March, 1997
This report is a project of the Roy Wilkins Center for Human Relations and Social Justice, Humphrey Institute of Public Affairs, University of Minnesota and a subcommittee of the Association of Metropolitan School Districts, Richfield, Minnesota. The Wilkins Center is directed by Samuel L. Myers, Jr. who served as the principal investigator of the project. The project was funded by the St. Louis Park Public Schools, Carol Johnson, Superintendent, through a grant from the Minnesota Department of Children, Families & Learning, Robert Wedl, Commissioner. General oversight for the project was provided by Irma McIntosh Coleman, Assistant Commissioner, Office of Teaching and Learning, Minnesota Department of Children, Families & Learning. Programming for the project and technical assistance was provided by: Dr. Ernest Davenport, Associate Professor, Department of Educational Psychology, University of Minnesota; David Heistad, Psychologist, Research, Evaluation and Assessment, Minneapolis Public Schools; Carol Hokenson, Supervisor, and John Comstock, Analyst/Technical Support, Data Management, Office of Information Technologies, State of Minnesota; Julie Buckel and Susan Hakomaki, Office of Information Technology, Academic and Distributed Computing Services, University of Minnesota. Research assistance was ably provided by Haijiang Kuang, Department of Educational Psychology, University of Minnesota. Additional research assistance was provided by Mary Lou Garza, Amanda Jameson, and Xuanmai Nguyen of the Wilkins Center, Humphrey Institute. Editorial assistance was provided by Dana Schroeder, Julia Blount, Cheryl Mandala, and Judith Leahy of the Wilkins Center.
The Problem

Recent press accounts document the poor achievement of many students on the Minnesota Basic Standards Test administered to all eighth graders in the state of Minnesota in 1996 and again in 1997. Reports of the 1996 results revealed wide racial and economic disparities in performance and significant disparities across jurisdictions. These accounts and the public commentary surrounding them suggest a level of certainty about the results that simply does not exist.

For example, there are reports that African American children do less well than white children in particular jurisdictions but that the central explanation for the poor performance of those children is that they are poor. While this is a plausible explanation, this conclusion is not the result of a rigorous econometric analysis.

Media accounts also suggest that the level of funding and expenditures for school districts is unrelated to the test-score performances. This conclusion seems to be reached by identification of some small school districts that performed well and some large school districts that performed poorly. The published accounts, however, do not reveal how this conclusion was reached or what methodology was used to reach it.

Policymakers and the constituencies they serve need the proper documentation and statistical analysis of the data before they can craft rational strategies for improving test scores among Minnesota students. This report provides the first comprehensive empirical analysis of the 1996 test score data and will assist local school superintendents, state officials and the public at large in evaluating the options available for reducing racial gaps in test scores and improving overall delivery of services in the state's schools.

Background

It is well-known that African Americans have consistently scored less well than whites on standardized tests (Coleman et al., 1966; Crossland, 1971; Fischer et al., 1996; Orfield, 1996; Scheirer, 1991). There is also evidence of lower test scores among Native Americans (Chavers and Locke, 1989) and Hispanics (Orfield, 1996; Veir, 1993), as well as mixed evidence concerning racial differences in test scores among Asian Americans (Orfield, 1996; College Board, 1988; Fischer et al., 1996; Coleman, 1966).

Various explanations for these discrepancies include differences in family background (Brooks-Gunn, 1996; Ceci, 1990; Currie and Thomas, 1995; Johnson, 1993; Medley
and Quirk, 1974; Venezky et al., 1987), economic resources (Brooks-Gunn; Crane, 1995; Fischer et al., 1996; Johnson, 1993), school resources (Jones et al., 1976; Maedaus et al., 1980), neighborhood poverty (Menacker, 1990; Yancey et al., 1995), teacher qualifications (Ehrenberg and Brewer, 1995; Ferguson, 1991), and academic background (Johnson, 1993). Others have attributed the gap in test scores to innate cognitive ability (Herrnstein and Murray, 1994), although there is great dispute over the size of the effects of cognitive ability on achievement (e.g., Johnson, 1995), as well as the causes of those effects. Some of the current literature on achievement tests shows persistent racial gaps even when controls are made for variables such as family background and socioeconomic status (Fischer et al., 1996).

The Minnesota Basic Standards Test is part of a nationwide policy movement towards basic skills testing of school children. The test grew out of the Trial State Assessment program, implemented as part of the National Education Goals, which were adopted in 1990. This program entails a voluntary assessment of eighth-grade achievement in mathematics and reading, by which states can compare their progress over time (NAEP, 1991). By the end of the 1980s, many had statewide basic skills tests (ETS, 1990). Such tests attempt to measure reading, math, and, more recently, writing skills (ETS, 1996), usually among children in grades two, four, eight and 12, to assure that high school graduates possess a uniform set of math and reading skills.

By the year 2000, students in Minnesota will be required to pass the basic-skills tests in math and reading in order to receive a high school diploma (Smith, 1996); in 2001, the writing test will be added as well. In the preliminary test given last year, only 63 percent of all eighth graders passed the reading section and only 76 percent passed the math test. Nearly one-third of the sample of students given the writing test failed it. In Minneapolis, only 42 percent passed the math test and 37 percent passed the reading test (Hotakainen, 1996). Given the controversy surrounding bias in standardized tests, it is crucial that we examine these test results with a critical eye.

The Questions

Our research focuses on the following questions:

- What is the effect of individual student characteristics on test scores?
- What is the effect of school characteristics on test scores?
- Can the racial gap in test scores be explained by individual student characteristics and/or school characteristics?
- Does the residual effect of race diminish further once account is taken of census
tract variables?

- What factors best predict high math and/or reading scores?
- What factors best predict low math and/or reading scores?
- Do minority high achievers have the same characteristics as nonminority high achievers?
- Do minority low achievers have the same characteristics as nonminority low achievers?
- Does that answer differ for different minority groups?
- If minority students had the same observed characteristics as nonminority groups, how much of the racial gap in test scores would remain?
- How much of the racial difference in test scores can be attributed to difference in school characteristics vs. student characteristics?

These are interrelated questions that require specific modeling and data analyses. The following section outlines the general patterns in the data. We then sketch out the models used and the data available for estimating the models. Then we summarize the key findings that address the questions posed above.

**Overview of the Data**

The data for the analysis that follows come from a merger of the MARSS (Minnesota Automated Reporting Student System) and the 1996 eighth grade Minnesota Basic Standard Test scores. There are 52,079 eighth graders in the data set of test results analyzed.\(^1\) Of this total, 1.66 percent are American Indians, 2.57 percent are Asian Americans, 1.33 percent are Hispanics, 2.85 percent are African American, and 91.59 percent are Caucasians. Not all students in the data set took both the mathematics and the reading examinations. There were 51,161 students who took the math exam and 49,326 who took the reading exam, excluding those for whom there was missing

---

\(^1\) We excluded persons in subsequent analysis whose ages were greater than 16 or less than 11.
information or absence of identifying information to match the MARSS and test score data.
1. Average Test Scores

The average test score is the easiest measure of performance to understand and interpret. Much of the focus in this report will be placed on this measure, called the Mean Test Score (MTS). The results are expressed as the average percent correct on each portion of the examination. Whites scored about 80 percent correct on the math and 70 percent correct on reading. All minority group members did less well than whites on the math and reading portions of the examination. Asian Americans scored 73 and 62 percent correct; American Indians, 65 and 57 percent; Hispanics, 65 and 54 percent; and African Americans, 58 and 54 percent correct.

Within each examination, however, there were wide variations in test scores. The math exam, which consisted of 68 questions, included modules on whole numbers and fractions, percentages and ratios, number sense, estimation, measurement, tables and graphs, chance and data, and shapes and space. The reading examination included literal and inferential items.

The whites scored 85 percent, 70 percent, 81 percent, 70 percent, 81 percent, 87 percent, 76 percent and 88 percent correct on the eight modules of the math exam. This means that they did much better on the whole numbers and fractions, tables and graphs, and shapes and space modules than they did overall. This also means that they performed less well on the percentages and ratios, estimation, and chance and data modules than on other parts of the examination.

African Americans, in contrast, scored 62 percent, 44 percent, 56 percent, 52 percent, 63 percent, 67 percent, 55 percent and 69 percent on the eight modules of the math exam. African Americans, therefore, performed very poorly on the percentages and ratios module but did much better on the tables and graphs and shapes and space portions of the examination. Their performance on the estimation and chance and data modules was not far below their overall average.

Asian Americans, American Indians, and Hispanics also scored higher on the shapes and space and tables and graphs sections than on the other sections of the math exam. They performed least well on estimation and chance and data.

The differences between the American Indian and Hispanic scores on various parts of the math exam are not statistically significant. These scores, however, generally are lower than the Asian and white scores and higher than the black scores. All of the differences in scores both within each math module and for the total math exam between nonwhites and whites are statistically significant at the five percent level.
There are also highly statistically significant disparities in reading test scores between nonwhites and whites. The reading scores for the literal module of the exam are 71 percent correct for whites; 63 percent correct for Asians; 58 percent correct for American Indians; 54 percent correct for Hispanics and 50 percent correct for African Americans. The reading scores for the inferential portion of the exam are: 68 percent correct for whites, 61 percent correct for Asians, 56 percent correct for American Indians, 55 percent correct for Hispanics, and 51 percent correct for blacks. This means that although whites, Asians, and American Indians scored slightly lower on the inferential portion of the reading exam, blacks and Hispanics performed slightly better on that part than on the literal part.

Thus, overall, there are wide racial disparities in the performance on the 1996 examination. The overall average math scores were below the pass mark of 70 percent correct for American Indians, Hispanics, and African Americans, while the overall average reading scores were below the pass mark for all nonwhite eighth graders taking the exam.

2. Distribution of Test Scores

Figures 1 - 12 plot the distribution of test scores for all eighth graders taking the math and reading examinations in the spring of 1996. One can determine from these plots the percentages of students who are doing very well (scoring 90 percent or better), those who are passing (scoring above 70 percent), and those who are doing very poorly (scoring 30 percent or lower).

**Who is Doing Well?**

Figure 1 shows that about 11 percent received scores of 90 or above on the math exam. However, only 0.8 percent of African Americans and 1.9 percent of American Indians scored 90 or above on math, compared with 3.6 percent of Hispanics, 8.4 percent of Asians, and 11.6 percent of whites.

Figure 7 shows that about eight percent of all test takers scored 90 or above on the reading exam. As in math, African Americans had the lowest share of top scorers on the reading test with only 1.3 percent receiving scores of 90 or above. Only 2.13 percent of American Indians and 2.11 percent of Hispanics scored 90 or above on reading, considerably below the 5.78 percent of Asians and 8.11 percent of whites scoring at that level.

Thus, at the top of the distribution, whites dominate; blacks and American Indians
seriously lag behind. While there were 5,450 whites in the state who received math scores of 90 or above, there were only 13 African Americans and 16 American Indians in the entire state who received such top scores.

**Who is Passing?**

Overall, 28.41 percent of students failed to earn a passing grade of 70 or better on the math portion of the exam; 44.83 percent failed the reading portion of the exam. In other words, about 72 percent of test takers received a passing score on the math exam; about 55 percent of test takers scored over 70 on the reading portion of the examination.

Among whites, 74 percent passed the math exam and 57 percent passed the reading exam. Among minorities, much larger fractions failed. Almost 60 percent of American Indians failed math and almost 68 percent failed reading. Nearly 60 percent of Hispanics failed math and 70 percent flunked the reading exam. Seventy-five percent of African Americans received scores below passing on the math exam, while an incredible 79 percent were below passing on the reading exam. For Asians, passing rates were higher, at 59 percent on the math exam and 44 percent on reading.

Thus, there were many minority eighth graders who failed to score at least a 70 on the math and reading exams, even though the majority of whites and Asians passed the math exam and the majority of whites passed the reading exam. Put differently, while most whites passed, most minorities flunked.

**Who is Doing Poorly?**

On a multiple choice examination it is frequently possible to answer up to half the questions merely by guessing. It is almost impossible to score less than 30 percent correct, if one simply chooses the answers randomly. From Figures 1 and 7 we see that only 7.84 percent of test takers scored below 50 percent correct on the math exam, while only 1.17 percent scored less than 30 percent. Only 15 percent scored less than 50 percent on the reading exam and 3.4 percent scored less than 30 percent. The small percentages receiving scores below 30 percent confirm the view that it is possible to guess your way through about a third of the questions on the examination. Still, in the overall population, even at the cut-off of 50 percent correct, the vast majority of test takers were able to do better than simply guess.

Unfortunately, more than 11 times the percentage of African Americans compared to whites scored below 30 percent correct on the math exam (9.02 vs. 0.79 percent). Six times the percentage of American Indians compared to whites scored below 30 percent correct on the math exam (4.95 vs. 0.79 percent). On the reading test, the racial gaps
at the bottom were a little smaller. African Americans were more than six times as likely to score below 30 percent correct as whites (18.14 vs. 2.63 percent); American Indians and Hispanics were both about three and a half times as likely to score below 30 percent correct as whites (9.4 and 9.74 vs. 2.63 percent); Asians were about three times as likely to score at the bottom as whites (8.19 vs. 2.63 percent).

Again, then, we find that blacks and other minorities perform less well than whites.

3. Race and Gender Differences in Test Scores

Figures 13 and 14 show that among whites, girls perform slightly less well on math than boys, while the reverse is true on the reading exam. By way of contrast, Asian girls perform slightly better on the math exam than Asian boys and much better than the boys on the reading exam. The lowest performances of all came from African American males and females. The African American boys’ scores on the math exam are slightly higher than those of African American girls, while the boys’ reading scores are lower. Among other races the math scores are not significantly different between males and females, although females generally have higher reading scores than males.

4. Specific Components of the Examinations

Figures 15 - 26 plot the scores by race for specific components of the reading and math tests. The first part of the mathematics test is comprised of estimation. These questions require the student, for example, to estimate the height of a table or the distance traveled by a train or to perform similar exercises without the aid of a calculator.

Some analysts consider this part of the examination to be most difficult. Yet, the black-white gap in scores was only 18 points on this part of the exam, as opposed to an overall gap in excess of 22 points on the entire math exam.

Once the estimation section is completed--students have as much time as they like to work on this section--the book is closed and the remaining portion of the mathematics exam can be performed using a calculator. The modules tested for these remaining parts include problem-solving questions involving a) percentages and ratios and b) whole numbers and fractions. As Figure 18 shows, all students performed least well on problem solving involving percentages and ratios. But, the black-white gap is huge, at 26 percentage points. On other more straightforward sections, such as shapes and space and whole numbers and fractions, all students performed well, with American Indian, Hispanics, and African Americans nearly passing or passing these sections. Asians and whites uniformly passed these sections handily.
On the reading tests, nonfiction reading materials were selected to gauge both literal and inferential skills. There were only minor differences in the scores between those two sections, with whites and Asians performing slightly better on the literal than the inferential, but with little statistical difference in the black, American Indian, or Hispanic scores between the two sections.

5. Relative Representation in Top- and Bottom-Scoring Schools

We computed the average math and reading test scores of each school in the state. We then ranked the schools from the top to the bottom. We divided the list of ranked schools into five categories. The top quintile or top 20 percent were labeled A(top-ranked); the bottom quintile, or bottom 20 percent, were labeled A(bottom-ranked).@ Figures 27 and 30 show the racial composition attending the top-ranked and bottom-ranked schools respectively, relative to the racial composition of all test takers. Blacks, American Indians, and Hispanics are severely underrepresented in the top-ranked reading schools and blacks and American Indians are severely underrepresented among the top ranked math schools.

Asians and Hispanics are more than twice as likely to be found in the bottom-ranked math schools as they are to be found in the eighth grade population. American Indians are 3.2 times as likely to be found in the bottom-ranked math schools as in the general population and African Americans are 4.5 times more likely. The disparity is smaller in reading rankings. Blacks are only twice as likely to be found in the lowest-ranked reading schools as in the general population of eighth graders.

6. Accounting for Zero Scores

Analysis of the test scores shows that nontrivial numbers of minority students obtained zero scores on the examination. Moreover, the examination was administered throughout the state to seventh graders, ninth graders, and even students in high school. What are the impacts of our restriction of the analysis to nonzero scorers and to eighth graders? Figures 31-34 show that the effects are negligible. There are only small differences in the racial disparities in test scores when account is taken of the exclusion of zero-scorers or of non-eighth graders.

7. Racial Gaps at the Top and at the Bottom

Descriptive analysis in Figures 35-36 address the question of whether there are racial disparities in test scores within categories of schools. All but a tiny fraction of the Asian-white gap in reading or math test scores disappears at the top-ranked schools.
There are generally gaps of about eight to 16 percent between whites and the other races in the top-ranked schools, with reading gaps jumping to 28 percent for whites and Hispanics and 21 percent for whites and Indians. At the bottom-ranked schools, racial gaps tend to be consistently larger. White scores are 23 to 28 percent higher than black scores in the bottom-ranked schools; the gap is about 19 to 22 percent between whites and Hispanics, 10 to 20 percent between whites and Asians, and 13 to 19 percent between whites and American Indians. This suggests that while there are also noticeable disparities at the top, blacks and Hispanics perform much less well than whites at the bottom.

8. High Achievers vs. Low Achievers

Tables 1 and 2 provide the characteristics of high achievers and low achievers among white, black, Asian, Hispanic and American Indian eighth graders. High achievers are those who scored in the top 20 percent of test takers. Low achievers are those who scored in the bottom 20 percent of test takers. The tables reveal that there are many important racial differences among high and low achievers, as well as several notable similarities.

Key Differences

**High Achievers**

Attendance rates, age, racial composition and poverty of schools, and ranking of schools differed between many minority students and whites. For example, Asian attendance rates are considerably higher than white attendance rates among high achievers. More Asian high achievers are girls than among other races. Both Asian and black high achievers are more likely to attend schools with large proportions of minorities than are whites. Black reading high achievers attended schools with an average of 14 percent black, while white reading high achievers attended schools with only 1.9 percent black. Among math and reading high achievers, Asians attended schools with five percent Asians, but whites attended schools with less than half that percentage of Asian students. School poverty rates were considerably higher among American Indians than among whites.

American Indians who were high achievers on math and reading tests attended schools that were ranked considerably lower than those of whites. Top ranking blacks were also more likely to attend lower-ranked schools than whites. Whereas only five percent of whites who scored in the top 20 percent on the math exam attended schools ranked at the bottom 20 percent, more than 25 percent of blacks attended these low-ranked schools. Whereas 7.5 percent of whites who scored in the top 20 percent of the
reading exam attended lower-ranked math schools, 35 percent of top-ranked blacks attended these schools.

In summary, minorities who excelled did not necessarily attend the same schools or have the same backgrounds as whites who were top achievers. One glaring disparity is participation in gifted and talented programs. Whereas 23 percent of the top white math performers were enrolled in gifted and talented programs, only eight percent of American Indians were. The disparity is even larger for reading. A full 24 percent of white top performers on the reading exam were enrolled in gifted and talented programs, while only 6.9 percent of top-scoring American Indians were. There are no statistically significant differences in the representation of blacks and Hispanics in gifted and talented programs. Top-performer Asians are slightly more likely than top-performer whites to be enrolled in gifted and talented programs.

Low Achievers

Attendance rates are considerably lower for many minority group members scoring at the bottom of the exam than they are among whites at the bottom. Low achieving minorities also had more school changes and greater difficulty with English, were enrolled in schools with greater percentages of minority students and higher school poverty rates, and were located consistently at schools ranked at the bottom. Interestingly, however, white low achievers were more likely than black, Asian, Hispanic, or American Indian students to be enrolled in Special Education or Disabled programs.

Similarities

Among all races high achievers tended to change schools infrequently. The average number of school changes in the past five years for this group was only about 2.2 to 2.5, insignificantly different among the various racial groups. This means that an eighth grader has changed schools only from elementary to middle school. (A student who has attended two schools in five years would have recorded two school changes: leaving the elementary school and joining the middle school.) In the sample some students registered as many as 13 school changes in a five-year span. Among minority low achievers of all races, the school changes are much larger in number than among high achievers. The only other major similarity between white and minority high or low achievers is gender. Because of the small numbers of minority high achievers in the sample, the black-white, Hispanic-white and American Indian-white differences in gender are not statistically significant. Among low achievers, there are no differences in gender by race on the math scores (most low math achievers are girls) and only a slight black-white difference on the reading scores (most reading low achievers are
boys). On balance, then, there are many more differences than similarities between
the white and nonwhite high and low achievers.

Model Specification and Data

We consider four specific model specifications in the analysis below. These are: mean
test score model, high achiever model, low achiever model, and the pass rate model.
Each model is estimated for math and reading tests, for all eighth graders, and then
within the eighth grade for each racial group. The models are also estimated for a
partition of the eighth graders in the state who were not enrolled in Special Education,
Disabled Student, Limited English Proficiency or Gifted and Talented programs.

In the mean test score model, the dependent variable is the average percent correct on
the specified examination. This model is estimated using ordinary least squares
methods.

In the high achiever model, the dependent variable is the probability that the student
scored in the top quintile (or top 20 percent) of all test takers. This model is estimated
by maximum likelihood methods using a logistic specification of the probability of high
achiever.

In the low achiever model, the dependent variable is the probability that the student
scored in the bottom quintile (or the lowest 20 percent) of all test takers. This model is
also estimated by maximum likelihood methods using a logistic function.

In the pass rate model, the dependent variable is the probability that the student scored
70 or above on the specified examination. A logistic specification is used here and
maximum likelihood methods are employed to obtain the estimates.

In each model there are three sets of independent variables. These can be classified
as individual-specific, program-specific, and school-specific. The independent variables are:

**Individual-specific characteristics (X)**

- *Race* (black, Hispanic, American Indian, Asian; left out dummy is white)
- *Gender* (female; left out dummy is male)
- *Age* (in years or months)
- *Attendance rate* (percent of membership days attended)
- *Mobility* (number of school changes in past 5 years)
- *English not spoken at home*
Program-specific Characteristics (Y)

Disability/Special Education (enrolled in Special Education or Disability program)
Limited English Proficiency (Enrolled in LEP program)
Gifted and Talented (Enrolled in Gifted and Talented program)

School Specific Characteristics (Z)

School poverty (Percent of students receiving free/reduced price meals)
Percent black in school (percent of all test takers at school who are black)
Percent Asian in school (percent of all test takers at school who are Asian)
Percent Hispanic in school (percent of all test takers at school who are Hispanic)
Percent American Indian in school (percent of all test takers at school who are American Indian)
Top quintile school - math (a dummy variable indicating whether a school ranked among the top 20 percent of mean math scores)
Top quintile school - reading (a dummy variable indicating whether a school ranked among the top 20 percent of reading scores)
Bottom quintile school - math (a dummy variable indicating whether a school ranked among the bottom 20 percent of math scores)
Bottom quintile school - reading (a dummy variable indicating whether a school ranked among the bottom 20 percent of reading scores)

These variables are available for nearly all schools and school districts in the state. In addition, we explore the mean test score model with the addition of three additional variables within one jurisdiction. These are:

Individual poverty (a dummy variable indicating whether the student was eligible to receive free or reduced price lunch)
Neighborhood poverty (the percent of persons below the poverty line in the census tract where the student resides)
Early Childhood Development (results of the 2nd grade California Achievement Test, nationally normed with a mean of 50).

In summary, then, the four models are:

High Achiever Model

This is a logistic specification of probability that the score was in top quintile of students (top 20 percent).
The coefficients $\alpha$, $\beta$, and $\gamma$ have the interpretation of being the change in the log-odds ratio of being a high achiever as a result of an increment in X, Y or Z. When these coefficients are negative, the effect is to lower the odds in favor of being a high achiever. When the coefficients are positive, the effect is to raise the probability of being a high achiever.

**Low Achiever Model**

Similar to the high achiever model, this is a logistic specification of the probability that the student's score is in the bottom quintile of all students taking the examination.

Positive coefficients, $\delta$, $\varepsilon$, and $\zeta$, mean that factors contribute to higher probabilities of being low achievers; negative coefficients mean that factors contribute to lower probabilities of being low achievers.

**Pass Rate Model**

In this model, the dependent variable is 1 if the score is 70 percent or higher. This is the probability that a student "passed" the examination. The model is estimated by logistic regression:
Mean Test Score Model

Most of the analysis to follow, however, uses the simple regression model where the dependent variable is the percent correct on the examination.

The coefficients $a^k$, $b^k$, and $c^k$ have the interpretation of being the change in the test score (percent correct) within the $k$th racial group as a result of a one unit change in $X$, $Y$ or $Z$. In the case of dummy variables, the coefficient has the interpretation of being the difference between the test scores of the included group and the excluded group defined by the dummy variable. Thus, if the dummy variable is female, then the coefficient is understood to mean the higher (or lower) score that being female entails.

The specific test of the determinants of the racial gap in test scores is performed by first estimating the MTS for whites and then computing the test scores for the $k$th racial group using the white coefficients and the $k$th groups=characteristics. The result provides a measure of the portion of the racial gap that cannot be explained by the variables in the model. Specifically, the unexplained gap is:
where,

This decomposition of the gap in test scores is explained further below. An analogous method is used to decompose the pass rate, the low achiever probability, and the high achiever probability.

KEY FINDINGS

In what follows, we summarize the results of estimating the four models for each of the racial groups and for certain subsamples. The estimated coefficients, model test statistics, and related information are detailed in Appendix Tables A1-A16. From these detailed results we obtain the information needed to assess:

- Effects of school poverty vs. attendance
- Effects of racial segregation vs. program participation
- Determinants of racial gap (via a decomposition analysis)
- The impacts of individual poverty vs. early childhood development

The results of the Appendix Tables are then summarized in relation to the original questions posed on page two of this report.

# School Attendance vs. Poverty

Attending school regularly is the most consistent and reliable predictor of success on the math and reading examinations. Although there are statistically significant racial differences in the rates of school attendance, the impacts of school attendance are not too different among racial groups. Table 3 clearly shows that a one percentage point increase in school attendance rates will raise test scores by about one-third of a point. Increasing school attendance rates from 80 percent to 95 percent will increase test scores by almost five points.

While this effect may seem small, it is substantially larger than the putatively adverse impact of school poverty on test scores. First of all, the poverty effect is not consistently significant and, when it is, it often has the opposite from the expected
negative sign. Only among Hispanics (reading scores) is the school poverty effect negative and significant--and its size just offsets the size of the attendance effect. That is, a one percent increase in school poverty lowers Hispanic reading scores, but those scores can be increased by about the same amount by increasing attendance. Among other racial groups the impact of school poverty on mean test scores is so small as to be inconsequential.

# Racial Concentration vs. Program Effects

Table 4 demonstrates that program participation had far larger impacts on mean test scores than did racial concentration within the school. First--among all eighth graders--the table examines the effects of own-racial representation and black representation in the school as compared to the impacts of participation in Gifted and Talented or Special Education programs. The panel clearly shows that diversity pays. White scores are lower the higher the percentage of whites in the school. White scores are higher the higher the percentage of blacks in the school. There are no statistically significant impacts of Hispanic representation on Hispanic scores, although Hispanic students do better at schools with larger fractions of blacks. There is no statistically significant impact of black representation on Asian or American Indian scores, even though both groups do less well with higher concentrations of their own group. Blacks do less well with higher concentrations of their own group on the math and reading examinations among all eighth graders. But when the second panel is examined--excluding special program students--the adverse impact of segregation on black reading test scores vanishes. There remain no adverse impacts of own-representation among Asians or Hispanics in this restricted sample.

This undoubtedly arises because of the tremendous influence exerted by program participation. The incremental effect of participating in a Gifted and Talented program is larger among minorities than among whites. Although the adverse impacts of being in Special Education or Disabled programs are not obviously worse for minorities than whites, the positive scores for being in a Gifted and Talented program considerably offset the negative scores for being in the Special Education programs.

On balance, nevertheless, the table shows that the impacts of program participation far outweigh the impacts of school poverty.

# Determinants of the Racial Gaps

The method used to decompose the racial gap in test scores is the following. First, we estimate from the underlying equations for a) mean test score, b) high achiever, c) low achiever, and d) pass rate coefficients for white eighth graders. These coefficients
measure the impacts on white students that the individual, program, and school characteristics have on the outcomes of percent correct on the examination, the probability of being a high achiever, the probability of being a low achiever, and the probability of passing. Using these coefficients to predict the outcomes for minority students is equivalent to asking this question: what would the test scores, high and low achiever probabilities and pass rates be for minorities if they faced the same treatment as whites? Or, equivalently, What would the outcomes for minorities be if they realized the same returns to each individual, school, and program characteristic as whites?@

Computing this equal treatment outcome and subtracting it from the actual outcome for minorities yields a residual difference. This residual difference is the unexplained gap in test scores, high and low achiever rates, and pass rates. The ratio of this residual gap to the actual gap yields the percent of the actual gap that is unexplained.@

Tables 5-8 reveal that there are large unexplained gaps in each of the outcome measures between minorities and whites. Table 5 refers to the black-white gap. Table 6 refers to the Hispanic-white gap. Table 7 refers to the American Indian-white gap. And, Table 8 refers to the Asian-white gap. For all 8th graders, the unexplained gaps range from 2 to 83 percent for Asians; 46 to 67 percent for American Indians; 35 to 70 percent for Hispanics; and 60 to 74 percent for blacks.

In each instance, we have computed the unexplained percent of the gap for reading and for math and for all eighth graders as well as for the sample of eighth graders excluding Gifted and Talented, Limited English Proficiency, Disabled, and Special Education program students. The results generally show a higher unexplained gap when the sample is restricted to non-special program students and, for blacks and Hispanics, higher unexplained gaps in the Mean Test Score model (percent correct) than the pass rate model. For Asians and Hispanics, the unexplained percent is smaller for reading than it is for math. For American Indians and blacks, the

2Another way to understand this is to conceive of the actual gap as consisting of two parts. One part is explained by racial differences in characteristics. The other part is what is left over (the residual) once the explained portion is accounted for. If the gap is not accounted for by the independent variables in the model, then what remains is the gap explained by differences in the returns to those independent variables or equivalent differences in the coefficients. Since the computation described above provides a measure of the gap accounted for by differences in the coefficients, then it is understood to be the unexplained portion of the actual gap.
unexplained percent is higher for reading than it is for math in both the high achiever model and the low achiever model.

To get an idea of what these results mean, consider the last finding of higher unexplained gaps for reading than for math in the American Indian and black high and low achiever models. Table 7 shows that 22.4 percent of whites and 5.8 percent of American Indians were high achievers in math while 18.6 percent and 6.7 percent of American Indians were high achievers in reading. The resulting American Indian-white math and reading high achiever gaps are 16.6 and 11.9 percentage points. The math gap is larger than the reading gap for this outcome measure between American Indians and whites. If American Indians were treated like identically situated whites, the math gap would fall to 7.7 percentage points; the reading gap would decline to 9.4 percentage points. If the residual gap were to disappear, then all of the actual gap could be attributed to differences in characteristics. The residual gap is smaller than the actual gap, but not by much in the case of math scores (9.4 versus 11.9). The result is that the unexplained percent is larger in the case of math (79.27% = 9.4/11.9 x 100%) than in the case of reading (46% = 7.7/16.6 x 100%)

**Black-White Gaps**

Table 5 reveals that there are uniformly high unexplained gaps across all of the models. Whether the outcome is high achiever or low achiever rates, pass rates or the mean test score, and whether the test is reading or math, there are huge unexplained gaps. For example, there were less than 3 percent of African Americans who were high achievers among those not in special programs. There were 16 to 20 percent white high achievers on the reading and math examinations among non-special program participants. The actual black-white difference in probabilities was about 14 percentage points on the reading exam and 17 percentage points on the mathematics examination. The residual differences were 12 and 11 percentage points. In other words, 90 percent of the high achiever rates between blacks and whites cannot be explained on the reading examination and 67 percent cannot be explained on the mathematics examination. Put differently, had blacks been treated like whites, far more blacks would have placed in the top quintile than the observed numbers. The disparity cannot be explained by differences in individual characteristics, school characteristics--and certainly not program characteristics, since the sample here is restricted to non-program participants.

**Hispanic-White Gaps**

Table 6 shows that there are reading disparities between whites and Hispanics that can be explained to some degree by differences in individual characteristics, school characteristics, and program characteristics. We know from the raw data that far more
Hispanics are enrolled in Limited English Proficiency (LEP) programs than are whites; and that far more Hispanics than whites come from homes where English is not the primary spoken language. But these factors also adversely affect test scores, especially reading test scores. That is why, much of the reading gaps are explained by the model in the case of Hispanics.

For example, 63 percent of white eighth graders passed the reading test. Only 35 percent of Hispanics passed the test. There is a 29 percentage point gap. The residual gap, however, is only 10 percentage points. Thus, only about 35 percent (10/29 *100) remains unexplained by the model.

**American Indian-White Gaps**

The smallest unexplained gap for American Indians is the high achiever gap on the mathematics examination. American Indians—even those who perform at the top—are far less likely than whites to be in Gifted and Talented programs. This program is a consistently important determinant of high test scores. As a result, the model (Table 7) explains much of the actual gap of 17 percentage points between white and American Indian rates of high achievement. Even if American Indians had been treated the same as identically situated whites—meaning whites who are also not in Gifted and Talented programs, who have the same levels of poverty, the same school, individual and program characteristics—the high achievement of American Indians would not be greatly changed. Instead of having high achievement rates of less than 6 percent, they would have high achievement rates of 13 percent (.0577 + .0766), still considerably less than the white achievement rate of 22 percent.

The largest unexplained gap for American Indians is the average reading scores for students outside of special programs. Almost 82 percent of the reading disparity is unexplained. The mean score for whites was 74 percent correct; for American Indians, it was 64 percent. This 10 percentage point gap is largely unexplained because had American Indians been treated like equally situated whites, their scores would have been nearly the same as that for whites, namely 72 percent correct (=63.68+8.45).

**Asian-White Gaps**

There are vast differences in the reading and math disparities among Asian 8th graders. Much if not all of the Asian-white reading disparities can be explained by the models. Since dominant factors in these models are participation in LEP programs and non-English speakers at home, the higher rates among Asians than among whites of poor English skills translates into explanations for the gaps in reading scores.
Similar explanations do not exert themselves in the mathematics models. For example, in the high achiever model, there is a 6 point gap between whites and Asians. Whites exhibit probabilities of high achievement in math at rates of 22 percent, while Asians exhibit probabilities of high achievement at rates of 16 percent. Had Asians been treated like whites, their probabilities of high achievement would have been higher by 5 percentage points (or 21 percent) and thus all but a fraction of the gap is explained. Table 8 shows that 83 percent is unexplained.

# Effects of Individual Poverty and Early Childhood Development

It is not possible to measure directly all of the important determinants of test scores with the data available. However, it is possible to examine the impacts of certain additional variables on test scores within a single large school district that systematically collects valuable measures of individual poverty and early childhood development.

Table 9 examines the impact of individual poverty and an indicator of early childhood achievement in reading and math concepts on eighth grade test scores. Individual poverty is measured by free or reduced lunch (or meal) eligibility. The 1990 California Achievement Test (CAT) scores are a proxy of early childhood development. These tests, which measure mathematics concepts and reading comprehension, were given to second graders. The tests are nationally normed, requiring that measures of the impacts on eighth grade test scores be computed on a common metric. We have chosen to measure the effects by the computation of elasticities. This is the percentage change in the dependent variable as the result of a one percent change in the independent variable. The elasticity is computed as the product of the coefficient estimated from the linear regression model of math and reading test scores and the ratio of the cognitive ability and poverty variables to the average test scores.

As is evident from the table, the average minority student did less well than the average white student on the math and reading tests. They are poorer. And minority students performed considerably less well on the California achievement exam six years earlier. Of course, the sample by necessity is restricted to students who were in the system during the second grade and excludes those who transferred from other states or other school districts within the state after the second grade. The sample also excludes those who transferred out of the particular district, the state, or out of the public schools altogether. Thus, the sample is not necessarily representative of all eighth graders in the district nor certainly of students throughout the state.

Nevertheless, the results are particularly telling. The poverty impacts are not statistically significant for American Indians, African Americans, Asian Americans, or Latinos/Chicanos. The only statistically significant impacts of poverty are found among
whites and even for them the impacts are minuscule. A one percent increase in poverty incidence (as measured by free or reduced lunch eligibility) reduces test scores by less than two-hundredths of one percent.

Impacts of early childhood development are larger. A one percent increase in the second-grade test score on the California Achievement Test increases eighth-grade math and reading scores from about one-fifth of one percent to more than one-half of one percent. The largest elasticity is registered for Hispanic math scores. The black and white elasticities are about the same and remain in the same range on both the math and reading tests. The impact of a one percent increase in black or white second-grade test scores is to increase eighth-grade test scores by about one-third of one percent.

In summary, individual poverty does not appear to be the proximate cause of the low test scores of minority students. Even among whites, where individual poverty is found to be statistically significant, the quantitative impacts on lowering test scores are small.

Part of the racial gap in eighth grade test scores, nevertheless, could be the result of racial differences in second-grade test scores, which captures an important aspect of school readiness and preparation in reading and mathematics. Along with other factors included in the model, early childhood development helps to narrow the racial gap.

To demonstrate this, we estimate the model for the chosen jurisdiction with and without the CAT test scores, information unavailable in the larger state data set. First we compute the percent of the racial gap that is unexplained by the variables available in the full data set, plus individual and neighborhood poverty, estimated for the jurisdiction. Then, we compute the percent of the racial gaps explained by all of the variables in the full data set, including individual and neighborhood poverty and second-grade California Achievement Test scores. The difference in the percent unexplained is attributable to the addition of the CAT scores to the analysis.

Table 10 provides the results from these computations. The table shows the white test scores, the nonwhite scores, the difference between the nonwhite and white test scores—labeled the Actual gap—at the predicted nonwhite test scores, the residual gap, and the percent of the actual gap that is unexplained. The predicted nonwhite test scores are computed using the white regression coefficients and the nonwhite means of the independent variables. These predicted values have two identical interpretations: the test scores of nonwhites had they faced identical returns to each independent variable as whites did; or the test scores of whites if they had nonwhite characteristics.

The difference between the actual nonwhite test score and the predicted nonwhite test score yields a residual gap. If all of the difference between the nonwhite and white test scores were unexplained, the residual gap would be zero. The percent of the actual gap that is unexplained is attributable to the addition of the CAT scores to the analysis.
scores could be explained by racial differences in the independent variables--the nonwhite vs. white characteristics--then there would be no residual gap. If the difference between the nonwhite and white test scores cannot be explained at all by racial differences in the independent variables, the predicted nonwhite test score would equal the white test score and there would be a residual gap equal to the actual gap. The ratio of the residual gap to the actual gap, therefore, is a measure of the portion of the actual gap that is unexplained by the variables in the model. The table shows this percentage before and after accounting for the California Achievement Test scores. The percentage difference between the resulting measures of unexplained percents notes the degree to which our measure of early childhood development contributes to the racial gap in test scores in the eighth grade.

Before summarizing the results of Table 10, we first note that there are substantial racial differences in the math and reading test scores in this sample. The black-white gap in math test scores is 24 percentage points; the Hispanic-white gap is 19 percentage points; the American Indian-white gap is 17 percentage points. The Asian-white gap is smaller at 13 percentage points, but statistically significant, nonetheless. There are smaller gaps in the sample of students who were in the system from the second grade and, thus, for whom we have the reading scores. The black-white gap in reading test scores is 21 percentage points; the Asian-white gap is seven percentage points; and the American Indian-white gap is 16 percentage points. These gaps are computed for all eighth graders. The table clearly shows, however, that much of the racial gap remains unexplained by individual poverty, school poverty, neighborhood poverty, program participation, and other school and individual characteristics. The exceptions are reading scores among Asians and Hispanics and math scores among Asians.

What happens to these gaps when account is taken of early childhood development indicators? The black-white math unexplained gap diminishes by nearly 40 percent; the Hispanic-white math gap falls more than 60 percent to a meager two percentage point residual difference; and the American Indian-white gap narrows by 30 percent. The Asian gaps are not computed for math, because the model predicts that Asians would have had lower math scores based on their second grade CAT scores than their actual math scores.

For the reading scores, gaps also narrow for blacks, Asians, and American Indians. The percentage gap widens for Hispanics, even though the actual gap falls by less than one percentage point. Here the actual gap is smaller among persons in the system from the second grade, meaning that the small residual gap accounts for a larger portion of the actual gap.

In summary, then, the effect of controlling for early childhood development, as
measured by the second grade CAT scores, is to diminish the unexplained portion of the racial gaps in test scores. Even so, there remains a 30 to 40 percent unexplained gap among blacks, a 30 to 35 percent unexplained gap among American Indians; a 20 to nearly 60 percent unexplained gap among Hispanics and a 30 percent unexplained reading gap among Asians.

Thus, even accounting for early childhood development, individual, school and neighborhood poverty, racial concentration of schools and school, program, and individual characteristics, there still remains a nontrivial racial disparity that cannot be accounted for by the included variables.

# Summary of Results

We have estimated four different models to measure various aspects of student performance on standardized tests using the merged MARSS and Minnesota Basic Standards Eighth grade test results of 1996. The models of mean test scores, high achievers, low achievers, and pass rates were used to answer questions about the determinants of performance and the factors contributing to the wide racial gaps in mathematics and reading scores. The answers to the questions posed are the following:

- **What is the effect of individual student characteristics on average test scores?**

  Children of all races performed better when they had fewer school changes, had higher attendance rates, and were younger. White females consistently had higher reading scores than did white males, although white males performed better than white females on the mathematics examination. Among all groups except blacks, non-English speakers scored lower on the math and reading exams than did English speakers. Among blacks, students coming from households where English is not the primary language scored higher on the mathematics test than did other blacks.

- **Do program characteristics matter?**

  Participation in Limited English Proficiency or Special Education programs generally lowered test scores--often substantially. Participation in Gifted and Talented programs uniformly increased test scores, often by much more among minorities than among whites. On the mathematics examination, there is a substantial minority payoff by participating in Gifted and Talented programs. Despite this fact, some minority group members, such as American Indians, are severely underrepresented in these programs.
What are the effects of school characteristics on test scores?

Measures of school poverty do not exhibit consistently negative impacts on test scores, once other factors such as racial composition and ranking of schools are accounted for. Diversity pays for white students: their scores are higher in schools with larger percentages of African Americans, Latinos/Chicanos, and American Indians. There is some evidence of adverse own-group impacts: higher percentages of non-Hispanic minorities in a school tend to be associated with lower test scores within those groups. Higher percentages of whites in a school tend also to be associated with lower white test scores. These impacts of diversity or minority concentration, nevertheless, are quite small. It would take, for example, an increase of 50 percentage points in the percent of black students in a school to lower the mean reading test score of African American eighth graders by 7.5 points. Moreover, much of the adverse impact of own-group segregation disappears once one focuses on students outside of Special Education and related programs. School characteristics, therefore, exhibit much smaller impacts on test scores than do program characteristics.

Can the racial gaps in test scores be explained by individual, program, and/or school characteristics?

Whether the model analyzed is that of mean test scores, high achievers, low achievers, or pass rates, the answer is the same: individual, school, and program characteristics measured in the merged MARSS and Minnesota Eighth Grade Basic Skills Test Scores data sets for 1996 explain very little about the racial gaps in performance. Although the percent unexplained varies from group to group and between the reading and math exams, the conclusion is the same: there is a huge overall unexplained disparity in test scores in Minnesota.

Does the residual effect of race diminish further once account is taken of individual or neighborhood poverty?

In the limited jurisdiction where this question can be addressed, the unexplained racial gap in test scores is just a little smaller than that for the rest of the state once account is taken of individual and neighborhood poverty. Moreover, the marginal impacts of neighborhood poverty are negligible, and often statistically insignificant. The effects of individual poverty are only evident among white students. Even then the impacts are not substantial.

What factors best predict high/low achievers?

High achievement in both the reading and math exams is most consistently related to
participation in Gifted and Talented programs across all races. High achievement is not statistically related to school poverty, racial concentration, or even ranking of schools—except in the case of whites. High scoring minorities can be found readily in even the lowest ranked schools. The impact on minority high reading scores of being in a top ranked school is generally insignificant. Being in a top ranked school does improve Asian, Hispanic, and American Indian students math chances of high achievement, but this is not so among African Americans. The ranking of schools is statistically unrelated to the probability of high achievement on the math or reading examinations among African Americans.

Low achievement in both the reading and math exams is also related to program participation. Persons in Special Education and Disability programs often are more likely to be low achievers. There are mixed effects of school ranking and racial composition on low achievement. Among blacks for example, being in a top ranked math school lowers the probability of being a low achiever in either reading or math; being in a bottom ranked reading school raises the probability of being a low achiever in reading; and, being in a school with a larger concentration of African Americans raises the probability of low achievement in reading, but not in math.

- **Do minority low/high achievers have the same characteristics as non-minority low/high achievers?**

There are far more differences between minority low/high achievers and non-minority low/high achievers than there are similarities. These differences involve attendance, school changes—especially among low achievers—racial composition of schools, English language proficiency, and school poverty.

- **Does the answer differ for different racial groups?**

Despite the many differences in characteristics across the racial groups and despite the often substantial differences between specific racial groups among high or low achievers, the more telling differences are between whites and minorities who are high achievers or low achievers. These differences, certainly, could help explain the sizeable racial gulf in high and low achievement.

- **If minority students had the same observed characteristics as non-minority groups, how much of the racial gap would remain?**

To answer this question, we computed measures of unexplained residuals which reflect the decomposition of the total gap in test scores between a) portions explained by the many racial differences in individual, school, and program characteristics, and b)
portions unexplained by these factors. We find large and substantial unexplained racial gaps in every outcome measure on both the reading and math examinations.

- **How much of the racial difference can be attributed to school characteristics vs. student characteristics?**

We examined the effects of individual poverty, neighborhood poverty, and second grade test scores on the measure of unexplained racial gaps in test scores. We find that accounting for individual poverty and neighborhood poverty does not reduce the substantially high unexplained residual gap in test scores. We do discover, however, that accounting for second grade test scores—a measure of early childhood education—does reduce the unexplained residual, often by 40 to 60 percent. Nonetheless, even with account being taken of early childhood development, there remain unexplained racial gaps in test scores of 20 to 40 percent on the mathematics exam and 32 to 57 percent on the reading exam.

**Conclusion**

We find non-trivial unexplained racial gaps in test scores in our various models. There are numerous possible reasons for these findings. Foremost among them is the possibility that we have failed to measure all of the factors that putatively explain test scores. Unexplored in this analysis are these important variables:

- Study habits, television watching, involvement in extracurricular activities, and other measures of time utilization
- Parental involvement, caretaker education, and assistance provided to children in preparation of home work assignments, reading at night, and similar nurturing activities
- Within school segregation, teacher race and education, teacher performance and motivation, administrative structure, and program effectiveness
- Costs and expenditures per student

It is possible that some or all of the unexplained racial gaps can be accounted for by these unmeasured or omitted variables. It is certainly possible that information about these factors could help improve our understanding of student performance and also enhance the policy making process designed to reduce or eliminate racial gaps in test scores. Some of these factors can be directly accounted for by additional data mergers and analysis. Information on teacher experience, race, education, and evaluations could be linked to student performance. Information on district or even school
expenditures in principle could be linked to student performance. Such efforts, however, would require some rethinking of how state data is collected and managed. Presently, the required information is not readily merged nor easily accessible for comparative analysis.

The call for more research is not the only implication of these findings, however. There are several clear and immediate initiatives that could in meaningful ways help improve test scores and reduce the racial gaps in performance on the math and reading examinations. These include:

! Greater efforts to increase school attendance, especially among minority students.

! Improved attempts to limit frequency of school changes and to stabilize student movements from elementary to middle school.

! Implementation of more and varied gifted and talented programs, particularly for American Indian students and others who are underserved by these programs.

! Careful assessment and evaluation of special education programs and their appropriateness for many minority students who might be better served in mainstream programs.

! A reexamination of English proficiency training and the impacts that targeted and non-targeted programs have on non-English speaking populations.

These efforts may have short-run implications for raising test scores and/or reducing the racial gaps in test scores. However, the finding of large reductions in the unexplained racial gap in test scores as a result of accounting for early childhood education suggests that more attention needs to be paid in the long run to improving second grade test scores. This will entail obtaining greater parental and caretaker involvement to improve reading and math readiness of minority children. This may require new and innovative programs designed to assist parents in understanding test taking methods and in learning how to monitor homework and how to assess children’s learning and performance. If greater burdens must be borne by parents to improve minority test scores, then we must determine what resources and knowledge parents need to do the job and to do it effectively. It does little good to say that parents must ultimately be responsible for their children’s math and reading readiness and then not to instruct or assist parents in fulfilling this responsibility.
References


