

RESEARCH BRIEF

Department of Research Services

FCAT Performance and the Achievement Gap

Introduction

Making Overall Comparisons

Since the beginning of the use of the FCAT by the State of Florida to evaluate the performance of schools, people have been interested in comparing the achievement levels of students in our school district with those of students throughout the state. On the whole, the district does not compare favorably. However, such comparisons rarely take into consideration the demographic differences between the district student population and the student population of the state.

Purpose of this Paper

In this paper the comparisons between the district and the state will be made within the subcategories of the major racial/ethnic groups. This categorization is regarded as especially appropriate because of considerable differences in performance levels between ethnic groups throughout the state. When the comparisons are thus constrained, the disparities between district and state performance are generally seen to disappear. The idea of adjusting for ethnic distributions is then extended to individual school considerations.

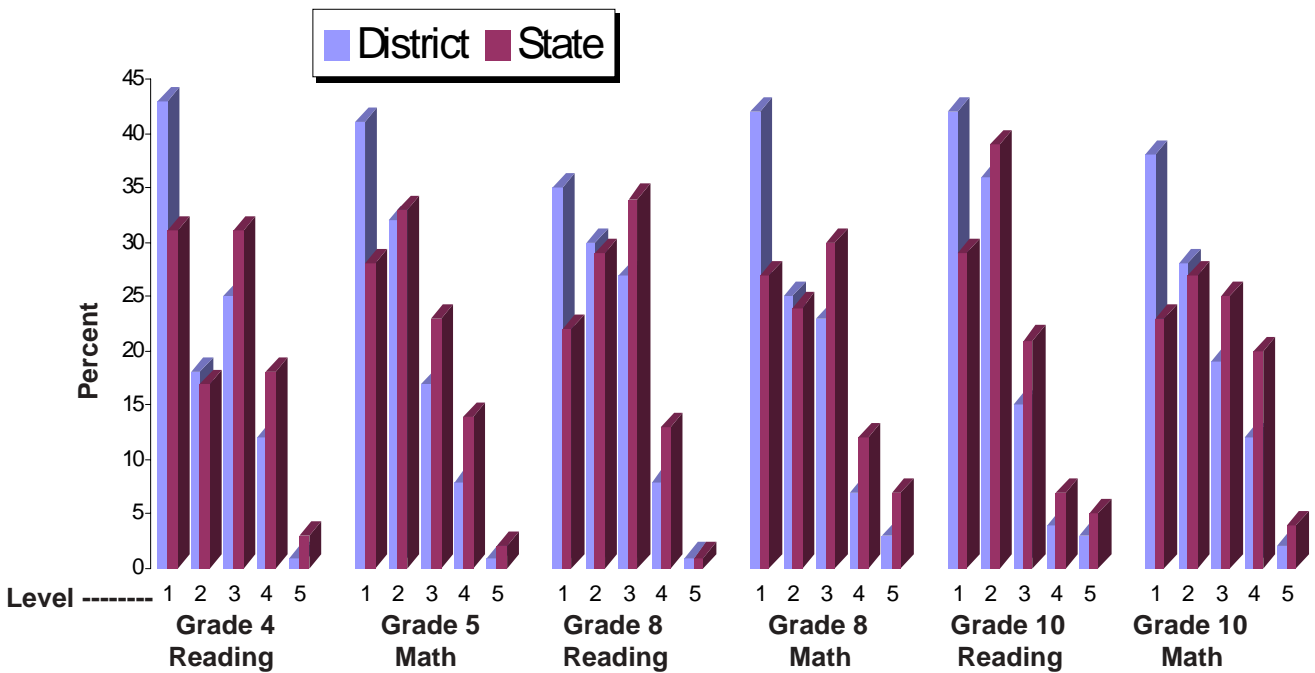
Technical Note: The Florida Comprehensive Assessment Test (FCAT) was originally designed as a criterion-referenced test which would typically use as its interpretive frame of reference a specified content domain rather than a specified population of persons. In practice, a normative framework is implicit in all testing, regardless of how scores are expressed. For example, by setting minimum competency standards on the FCAT, the State presupposes information about what other persons have done in similar situations. Assigning grades to schools is an obvious use of normative perspective and it is in this sense that we consider comparisons between populations with respect to FCAT performance.

Overall District -- State Comparison

Comparing Performance Level Distributions

Perhaps the simplest assessment of District performance would involve comparisons of overall performance level distributions between the District and the State. The graph below presents the percentages of students scoring in each of the five performance levels for each subject area of the 1999 administration of the FCAT.

Technical Note: The total score that a student could achieve on each subject area on the FCAT ranges from 100 to 500. These scaled scores are then converted to one of five possible performance levels based on score ranges designated for each subject area and grade level. The performance levels are interpreted as reflecting the degree of success the student has with the content of the Sunshine State Standards, with Level 1 representing the least achievement and Level 5 representing the greatest achievement. For reasons of simplicity and ease of cross-test comparisons, this paper will only consider test scores reported by performance level.



Interpreting the Differences

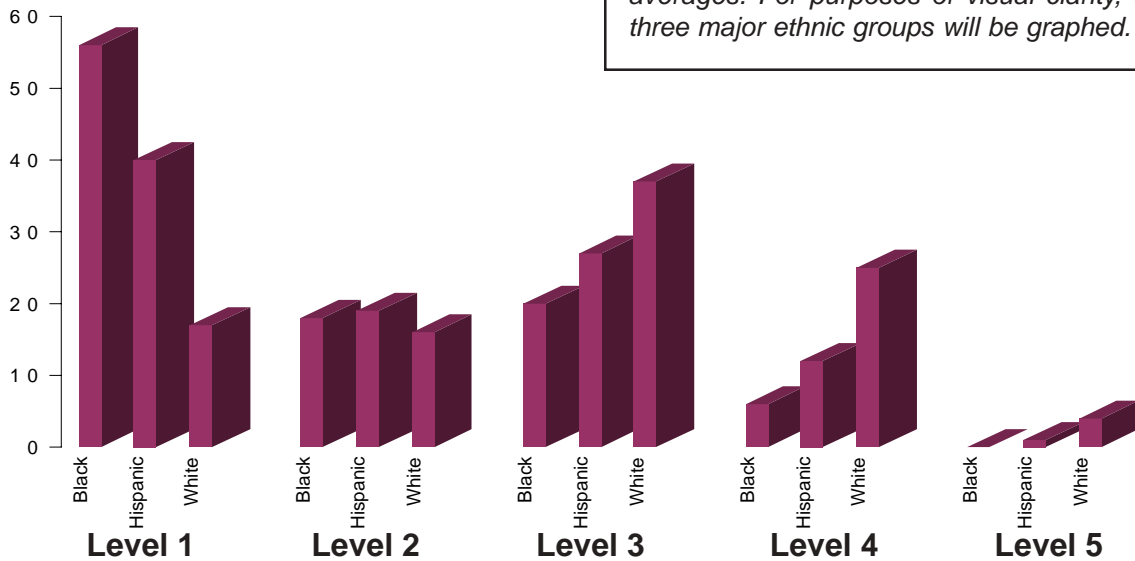
In general, the District exhibits a much higher percentage of students in Level 1 on all tests than does the State and a much lower percentage of students in Levels 3, 4, and 5. This paints a picture of below average performance by the students in the District.

The Statewide Achievement Gap

Disparity in Performance

One of the other phenomena that deserves attention is the “achievement gap” in FCAT performance. In this paper, the achievement gap refers to the disparity in performance that we observe when the test scores are disaggregated by race and ethnicity. As an example, the graph below depicts the statewide student percentages scoring in each of the five performance levels for the three major racial/ethnic groups in the Grade 4 Reading FCAT.

Technical Note: Of course, there are more than three racial/ethnic groups recognized in the student body of the district. The collective numbers for all those but the three major groups are comparatively so small that they have a negligible effect on district averages. For purposes of visual clarity, only the three major ethnic groups will be graphed.



Ethnic Differences

One can easily see the rather dramatic differences in ethnic distributions across the performance levels. The level with the greatest proportion for Blacks and Hispanics is the lowest level, Level 1. Whites are more evenly distributed across the levels with the greatest concentration being in Level 3. This kind of disparity is not of recent origin, nor is it new to the FCAT. This kind of phenomenon has been observed in other testing contexts in the past and has been a focus of concern in our district for many years.

Interpreting the Achievement Gap

It is important to avoid oversimplified interpretations of the achievement gap. There is, of course, nothing inherent in the ethnicity of students that *causes* test performance levels. However, these racial/ethnic categories are *related* to performance differences. According to a recent report by the

National Task Force on Minority High Achievement (*Reaching the Top*, the College Board, 1999):

“Since the 1960s, one of the big advances has been in our understanding of the sources of differences in academic achievement among racial and ethnic groups. With regard to expanding and improving efforts to increase the number of top underrepresented minority students, some of the most valuable insights concern five factors found to be strongly associated with student educational outcomes: 1) economic circumstances; 2) level of parents’ education; 3) racial and ethnic prejudice and discrimination; 4) cultural attributes of the home, community, and school; and 5) quality, amount, and uses of school resources.”

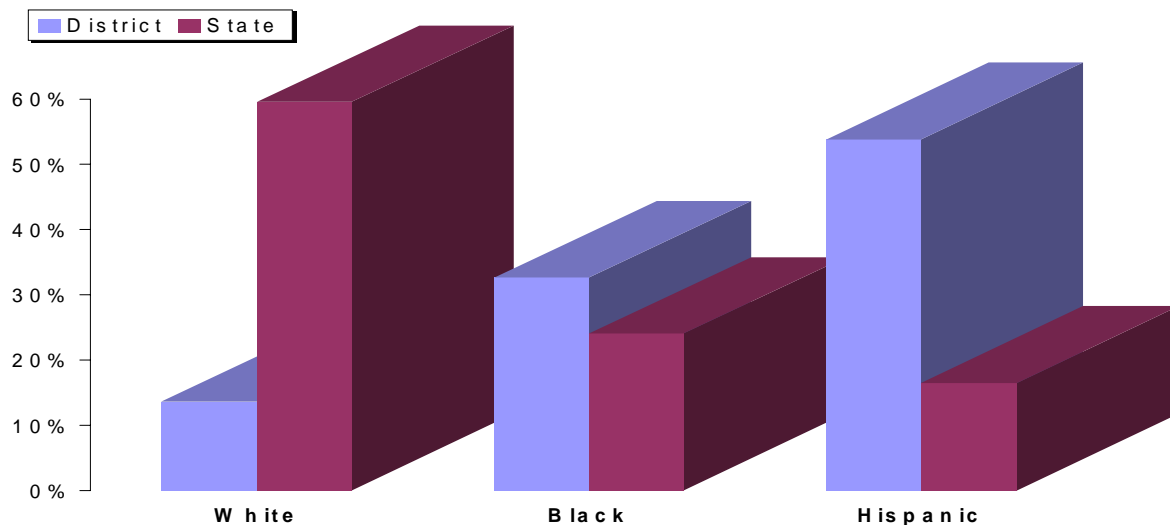
Acknowledging Differences Exist

There can be little doubt that our school district has both strong moral and practical interests in taking actions to ensure that minority groups significantly increase their academic achievement. Attempts to understand the underlying forces and direct energy and resources toward resolving the disparities have been long-standing goals of the district. As important as this is, it is not the topic of this paper. It is sufficient for this discussion to acknowledge that the achievement gap exists and to try to interpret our district performance with regard to that awareness.

Comparing Ethnic Distributions

Major Differences

Our district has a different racial/ethnic composition than the state as a whole. This is true in many ways, but none as dramatic as simple comparisons among the major ethnic groups defined by the state. The graph below displays these differences.



Taking Differences into Consideration

As can be observed, the state has a considerably larger percentage of White, Non-Hispanic students, a somewhat smaller percentage of Black students, and a considerably smaller percentage of Hispanic students. Sometimes when one is comparing some overall variable between the State and the District, the variable in question may actually be a composite of three different distributions, divisible on the basis of ethnicity. If such is the case, the composite will be weighted differently for the state than for the district because of the discrepant proportions of the ethnic groups. This is exactly the case in FCAT performance com-

Comparing Within Ethnicity

Equivalent Performance

The following two pages present graphic comparisons between the state and the district on FCAT Reading and Mathematics. The special feature of these graphs is the subcategorization of the scores into the three major ethnic groups. There is a remarkable degree of similarity between the district and the state in the within-group percentages and across-group patterns. **On the basis of this type of comparison, one would have to conclude that the district is performing almost exactly at average levels in the state, on par with typical student achievement.**

Questioning the Comparisons

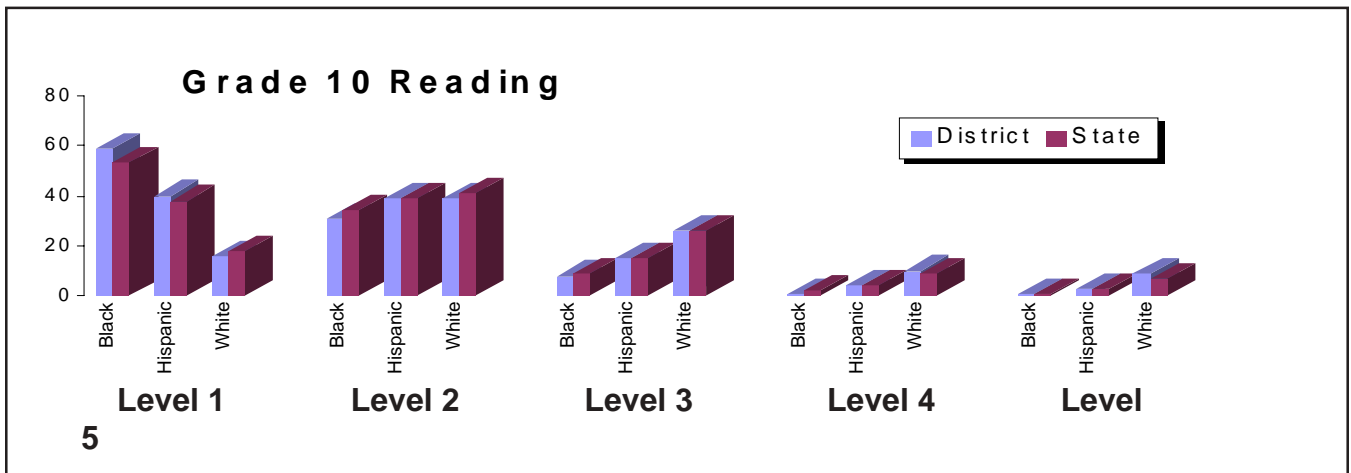
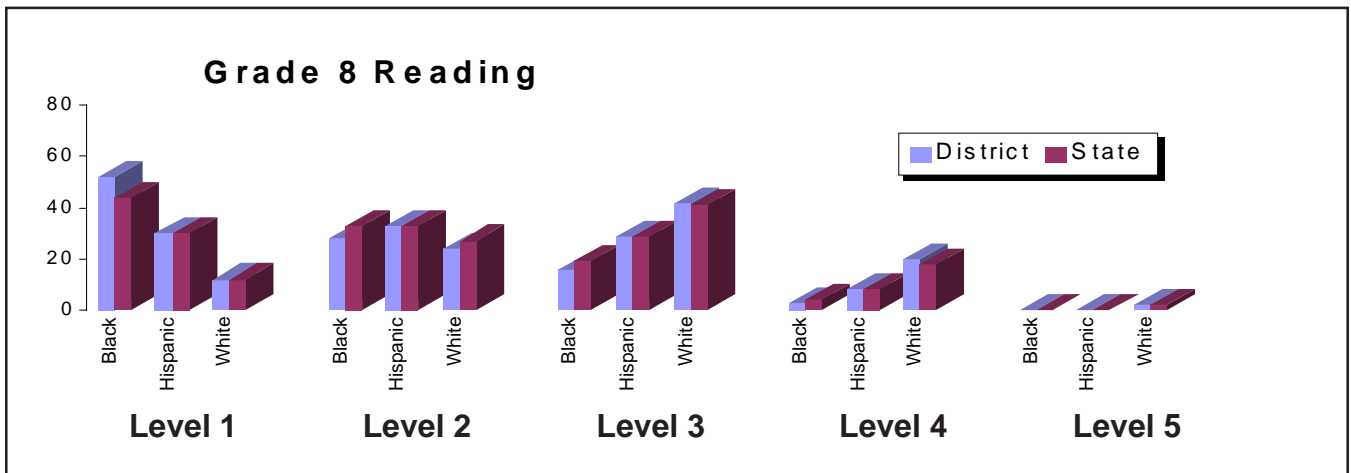
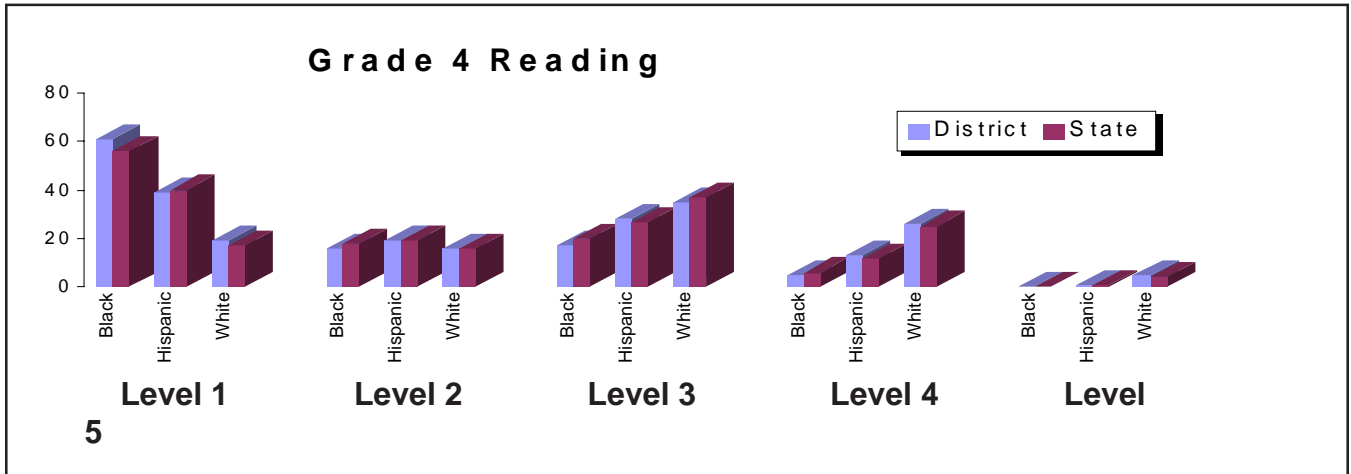
The question will be asked whether it is proper for us to make appraisals of the general achievement level of the district within the categorization of students by ethnicity. In doing so, are we presuming a lower standard for one ethnic group opposed to another? Is such a comparison an attempt to statistically obfuscate the more central issue of overall deficiency? The answers to these questions depend, in part, on the nature of the variables and the purposes of assessment. However, no presumptions of standards are made in this analysis. And overall deficiencies may be more apparent than real.

Consistent Patterns

The patterns of performance levels for each ethnic category are systematically consistent. Whereas there is a good deal of variety within any one ethnic group, the proportions scoring at each performance level remain stable across different contexts. This would, no doubt, remain true for other school districts. To a large extent, differences in FCAT achievement levels among districts is largely a function of differences in the racial/ethnic proportions for the districts.

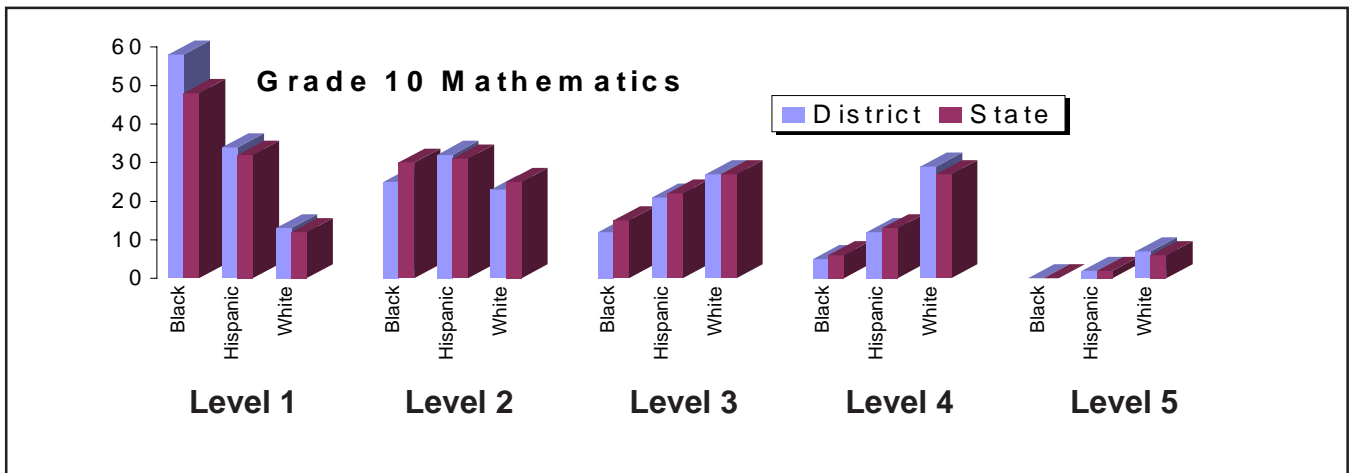
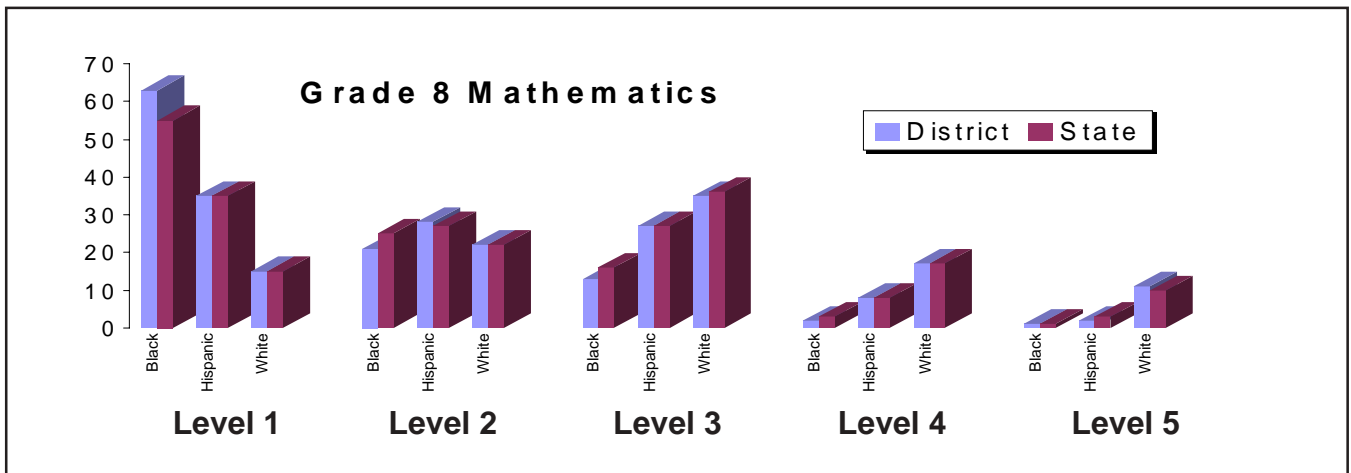
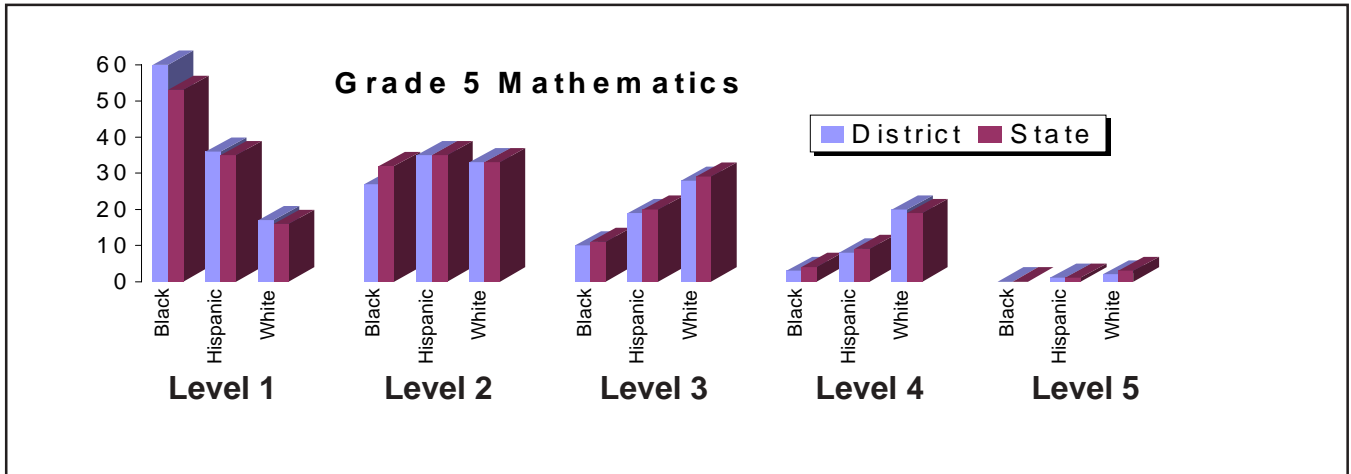
FCAT 1999 Reading

Comparing District to State by Ethnicity



FCAT 1999 Mathematics

Comparing District to State by Ethnicity



Predicting School Grades

Assigning Hypothetical Scores

One way to test the strength of the influence of the achievement gap at the school level involves assigning students hypothetical scores. Our purpose is to see how many schools could be successfully predicted to fail based only on average performance of each ethnic group. We can ascribe scores to students of one particular ethnic group such that the distribution of scores at a school matches the statewide distribution for that ethnic group. Thus, if 40 percent of Hispanics score at Performance Level 1 across the state, we assign scores to the Hispanic students at a given school so that 40 percent of those students score a Level 1. And if 19 percent of the Hispanics score at Level 2 statewide, we assign scores so that 19 percent of the Hispanic students at the school score at Level 2. This process is continued for all performance levels for all ethnic groups.

Technical Note: Actually, this process is easier than it might at first seem. It involves assigning proportions of proportions of scores -- an elementary operation with which few researchers would argue.

Defining Average Performance

The result of the assignment of scores is that each real school, with a real student body of given ethnic proportions is given a hypothetical set of scores such that the ethnic group performance exactly matches state averages. Thus, differences between overall school performance distributions for our hypothetical scores are attributable only to differences in ethnic proportions.

Criterion for Success/Failure

The State has set certain minimum criteria for school performance. Regarding the Reading and Mathematics FCAT tests, any school that has 60 percent or more of its students scoring at Level 2 and above will, at least, receive a grade higher than F. So, for convenience, we might say that a *school* has failed an FCAT test if it has not realized the 60 percent criterion.

Predicting Failure Based on Ethnic Proportions

Using this interpretation as a yardstick, we can see how many of our schools with hypothetical scores would be predicted to fail each test. We can then compare the predictions with the actual performance of the schools.

Successful Prediction

Having conducted this test and having compared predicted school failures to actual school failures, we find 80 percent of our schools can be predicted successfully. That is, **for each FCAT Reading and Mathematics test, at each level of Elementary, Middle and Senior High, failure of the school on the test can be successfully predicted 80 percent of the time based solely on the ethnic proportions in the school.** Apparently, the strength of the achievement gap is influential at the individual school level.

Conclusion

Individual vs. Group Expectations

The pattern of achievement for any ethnic group in no way restricts the potential for individual accomplishment. There are numerous examples of students achieving well beyond the average expectancy based on their ethnic group performance. In this sense, it is more than mere rhetoric to claim that all students are capable of high achievement. One may have justifiable optimism that the achievement gap will one day disappear.

Unequal Goals

However, for the time being, the differences in performance profiles for the ethnic groupings do exist. And imbalances in ethnic proportions can put districts and schools on unequal footing. Raising expectations can be an effective method of promoting higher achievement, but raising the bar to a common cutoff point can represent an inch increase to some and a foot increase to others. **How can it be proper to speak of “failing schools” when the designation of failure is predictable from the ethnic distribution in the school without reference to the actual assessment of the students?**

Proper Focus

The consistency and predictability of ethnic group differences suggests that, whatever the underlying causes, they appear to be generally resistant to influence by individual teachers. The real benefit from making comparisons within ethnic group categories is to redirect the focus of attention. The question is not “Why is this school not performing as well as it could,” but “Why is this ethnic group not performing as well as it could?” Investigating schools in which achievement is high *in spite of* ethnicity-related suppressors may yield useful tools for helping all students reach their full academic potential.