oppaga
OFFICE OF PROGRAM POLICY ANALYSIS
& GOVERNMENT ACCOUNTABILITY

April 2005 **Report No. 05-22**

Charter School Review Technical Report

This technical report provides additional details about the information and analyses presented in *Charter School Performance Similar to Other Public Schools; Accountability Needs Improvement*, Report No. 05-21. These reports assess the performance of students attending charter schools in the 2003-04 school year. We analyzed six years of Florida Comprehensive Assessment Test (FCAT) results to assess students' proficiency in math and reading before they entered a charter school, a year after they had attended their charter school, and their annual learning gains while in the charter school.

Our major findings were that, on average, charter school students were slightly behind academically before entering charter schools compared to students in traditional public schools; that charter school students were thus more likely than other public school students to not meet grade-level expectations; and the academic learning gains of charter school students were generally comparable to those of traditional public school students, although high school students at the lowest academic levels made slightly greater learning gains in charter schools than in traditional public schools. These findings are discussed in detail in our companion report.

This technical report provides information on

- the data sources used in our analyses,
- the methodology used to analyze whether students met grade-level expectations,
- the rationale and statistical models used to analyze student learning gains, and
- caveats and additional statistical analyses supporting the learning gains analyses.

Data Sources

We based our analysis of the demographic characteristics of charter and traditional public school students on the 2004-05 Survey Two fall data provided by the Florida Department of Education. We also analyzed student-level Florida Comprehensive Assessment Test (FCAT) data supplied by the department. This student-level data included all public school students in grades Pre-K through 12 during 2003-04. The FCAT data included the students' FCAT scores from 1998-99 through 2003-04.

Our analysis of students' FCAT performance used two related FCAT metrics: Florida's Sunshine State Standards (SSS) achievement levels and developmental scale scores (DSS). There are five FCAT SSS achievement levels. By statutory authority the Commissioner of Education has designated students at levels one and two as having limited success with the content of the Florida's Sunshine State Standards. Students at levels three and four are proficient, while students at level five are designated as advanced.

Each SSS achievement level corresponds to a range of developmental scale scores. FCAT developmental scale scores are vertically aligned allowing for the analysis of annual learning gains. The scale for reading developmental scale scores is from 86 to 3008. The scale for math developmental scale scores is from 375 to 2709.

Analysis of Students Meeting Grade-Level Expectations

We analyzed students' FCAT SSS achievement levels to determine if students were meeting or exceeding grade-level expectations. For 2003-04, we compared the percentage of charter and traditional school students that scored at achievement level 3 or above on the math and reading FCAT SSS (see Exhibit 1). Using information provided by the department, we converted students' FCAT developmental scale scores into FCAT SSS achievement levels. The classification of FCAT development scale scores into FCAT SSS is available on the Florida's Department of Education's website.

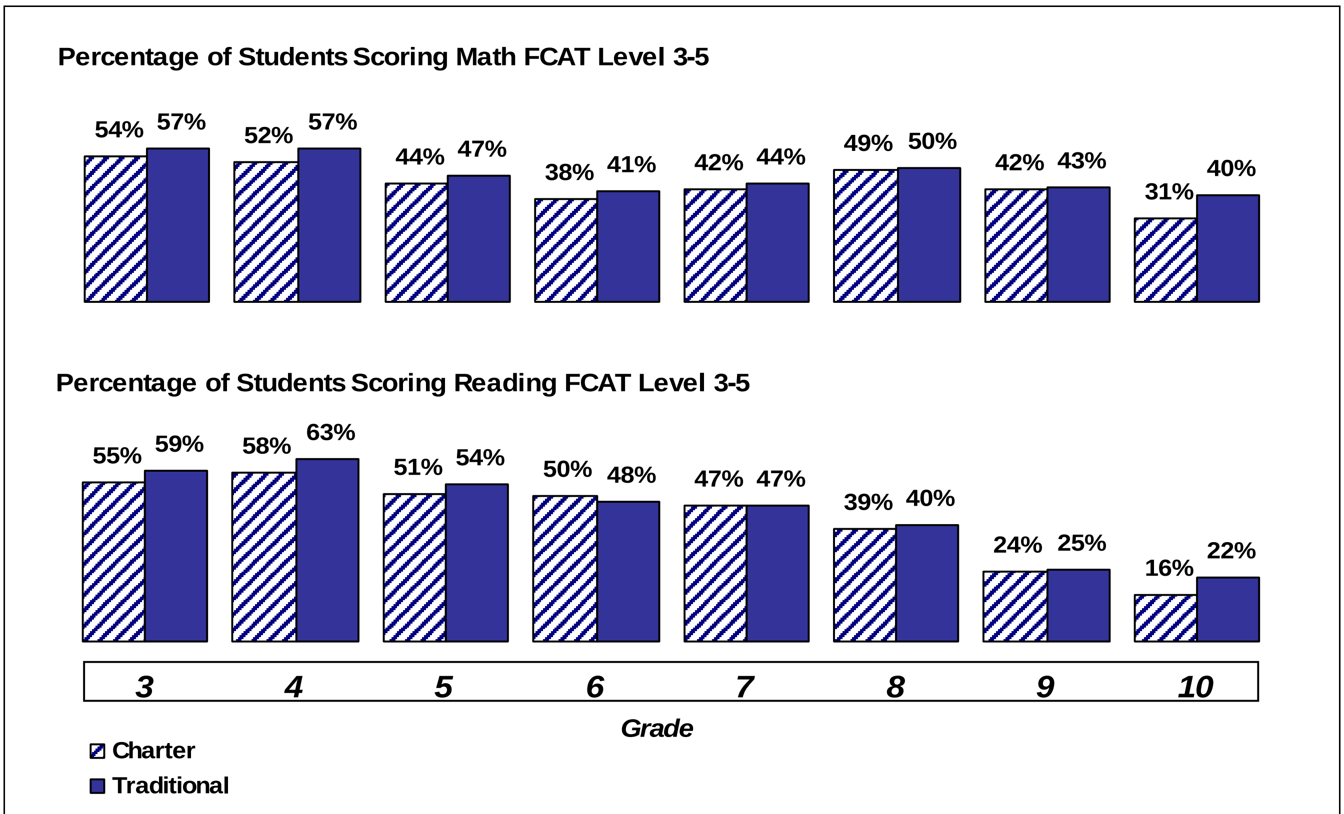
Analysis of Learning Gains

We used three-level hierarchical linear models (HLM) with FCAT developmental scale scores as the dependent variables to analyze whether the annual learning gains of charter school students were different from traditional public school students.

HLM Rationale

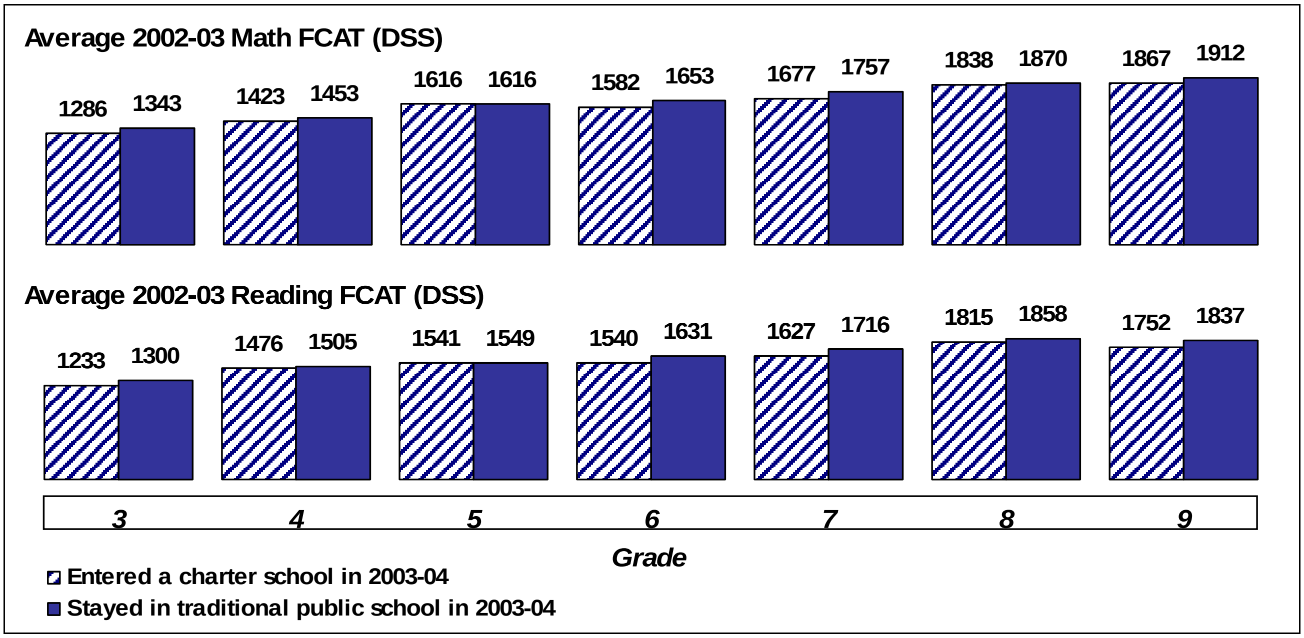
Snapshots do not account for students' previous academic performance. There is a risk of bias when comparing the academic performance of charter school students to traditional public school students. This risk can be reduced by comparing students' performance over time rather than by looking at only one year of data. The average charter school student is behind academically before entering a charter school (see Exhibit 2).

Exhibit 1
Charter School Students Were Less Likely to Meet Grade-Level Expectations in Math and Reading in 2003-04



Source: OPPAGA analysis of Department of Education data.

Exhibit 2
Students Entering Charter Schools Generally Had Lower FCAT Scores in Math and Reading the Previous Year Compared to Students Staying in Traditional Public Schools



Source: OPPAGA analysis of Department of Education data.

Because of this systematic difference and the fact that many charter school students have been in a charter school for only a short time, snapshots generally show charter school students not performing as well as traditional public school students. However, this perspective is flawed as charter school students may be making greater learning gains but have not yet caught up to traditional public school students. To correct for this, we accounted for each student's past performance and annual student learning gains by using multiple years of FCAT developmental scale scores. HLM is designed for longitudinal analysis of "nested" data and overcomes estimation problems of "nested" data with ordinary least squares regression models (Raudenbush and Bryk, 2002). [1]

HLM avoids misconstruing student and school "nested" effects. An advantage of HLM to compare charter school student performance to traditional school student performance is that it avoids limitations of non-hierarchical statistical models. Such models may misidentify differences due to a student characteristic with a charter school effect. Charter schools are schools of choice and may attract students that are different from those attending traditional public schools. For example, many charter students may come from homes where parental involvement in the student's education is higher than average. The possibility of incorrectly estimating student and school effects is greatly reduced by properly specifying each level in a multi-level model such as HLM. The hierarchical linear model that we developed controls for student characteristics and school characteristics in equations at separate levels. This reduces the risk that variation arising from student characteristics (such as parents who are actively involved in a student's education) is attributed to the school level or that variation arising from school characteristics such as charter status is attributed to the student level.

[1] Raudenbush and Bryk (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods* (2nd ed.). Sage Publications.

Hierarchical Linear Model Specification

We used three level HLM models in our analysis.

Level 1. The level 1 equation below depicts students' FCAT developmental scale scores (Y) as a function of time. The intercept P0 is the grand mean of students' developmental scale scores. P1 is the estimated annual learning gain. The test scores are "nested" within students, i.e., multiple years of scores for each student.

$$Y = P0 + [P1 \text{ H } (\text{TIME})] + E$$

Level 2. The level 2 equations depict how student characteristics affect developmental scale scores and annual learning gains. The level 2 equations below model the intercept and slope of the level 1 equation as a function of student demographic characteristics and a measure of each student's baseline developmental level prior to the developmental scale scores modeled at level 1. The formal specifications of the level 2 equations are as follows:

$$\begin{aligned} P0 = B00 &+ [B01 \text{ H } (\text{EXCEPTIONAL})] + [B02 \text{ H } (\text{GIFT})] \\ &+ [B03 \text{ H } (\text{LEP})] + [B04 \text{ H } (\text{LUNCH})] \\ &+ [B05 \text{ H } (\text{AFRICAN-AMERICAN})] \\ &+ [B06 \text{ H } (\text{FCAT STATUS t-1})] + R0 \end{aligned}$$

$$\begin{aligned} P1 = B10 &+ [B11 \text{ H } (\text{EXCEPTIONAL})] + [B12 \text{ H } (\text{GIFT})] \\ &+ [B13 \text{ H } (\text{LEP}) + [B14 \text{ H } (\text{LUNCH})] \\ &+ [B15 \text{ H } (\text{AFRICAN-AMERICAN})] \\ &+ [B16 \text{ H } (\text{FCAT STATUS t-1})] + R1 \end{aligned}$$

Level 3. At level 3, students are "nested" within schools. The level 3 equations below model the effect of school characteristics on the intercepts and slopes of the level 2 equations for P0 (the grand mean of students DSS scores) and P1 (students estimated annual learning gains). The level 3 equations test whether the grand mean of students' FCAT DSS scores is different for charter and traditional public school students and whether the learning gains of charter school students are different from traditional school students. We examined several different model specifications for the level 3 equations and selected an equation that accounted for the effect of charter on the level 2 parameters B00, B06, B10, and B16. These equations are listed below.

$$\begin{aligned} B00 &= G000 + [G001 \text{ H } (\text{CHARTER})] + U00 \\ B06 &= G060 + [G061 \text{ H } (\text{CHARTER})] \\ B10 &= G100 + [G101 \text{ H } (\text{CHARTER})] + U10 \\ B16 &= G160 + [G161 \text{ H } (\text{CHARTER})] \end{aligned}$$

Variable Descriptions. "TIME" at level 1 is a measure of the year in which a student took the FCAT. For example, level 1 cases that measure a student's first FCAT score have a value of "TIME" that is equal to 0. The value of time associated with a student's second FCAT score is 1.

We coded the level 2 student demographic characteristic variables a 1 or a 0. "Exceptional" is coded 1 for students with a primary learning exceptionality that was other than "gifted". "Gifted" is coded 1 if a student's primary exceptionality was gifted. "LEP" is coded 1 for students with limited English proficiency.

We coded students as 1 for "LUNCH" if at any time during the six-year period covered in our analysis they participated in the lunch program. Older students tend not to apply for the free lunch program even though they were eligible in earlier grades. Eligibility for the lunch program is commonly used to indicate a student's economic status.

"African-American" is coded a 1 for African-American students. The primary race and ethnic groups in Florida are African-Americans, Hispanics and whites. Our data did not distinguish between white Hispanics and black Hispanics as is commonly done. In deciding to only include "African-American" to measure a student's race or ethnicity we compared the correlations of "LEP" with a variable coded 1 if a student's ethnicity was Hispanic. The correlation between the variables "LEP" and "HISPANIC" was high. Including both variables in the HLM model did not substantially increase the reliability coefficient of the model. For this reason we included the variable LEP in the HLM model reported above but we did not include the variable Hispanic. "FCAT status t-1" is the student's baseline

FCAT developmental scale score measuring their initial developmental level.

We coded the level 3 variable "CHARTER" a 1 for schools that were a charter school in 2003-04, otherwise a "0".

Alternative specifications of students' baseline developmental scale score. We "centered" the baseline FCAT developmental scale score by the lowest student score at each school level (e.g., the lowest baseline developmental scale score for all high school students). We chose the lowest FCAT DSS score as the basis for comparison because the relationship between baseline score and learning gains is not linear, but is stronger at lower grade levels. Students at lower developmental levels, on average, achieve larger learning gains.

We also ran additional models in which we "grand mean centered" students' baseline score. This differs from centering by the lowest score. When grand mean centering, the basis for comparison is the average FCAT DSS rather than the lowest FCAT DSS. In our case where the differences between charter and traditional students were driven by low-performing students, the effect of grand mean centering was to inappropriately reduce performance differences. As a result the grand mean centered models did not show a statistically significant difference in the gains of charter and traditional school students.

Alternative specifications of school characteristics. The only school characteristic we included was the type of school, charter or traditional public school. We categorized the type of school a student attended as a charter or a traditional public school based on the type of school the student attended in 2003-04. We chose to include only charter status because we also examined models in which we specified school-level demographic variables in level 3, but these models did not substantially alter the statistical significance or size of the charter effect. For example, regardless of whether each school's percentage of students participating in the free and reduced lunch program was included or excluded at level 3, the estimates for the effects of charter status did not change substantially.

Case selection. The students in our analysis had a 2003-04 FCAT developmental scale score and at least two, but up to five prior consecutive years of FCAT developmental scale scores in their 2003-04 school. We considered using data for students for whom we did not have 2003-04 data, but for whom we did have data from prior years. We did not pursue this approach because we wanted to estimate the effect of schools that were charters in 2003-04, not charters that closed prior to 2003-04. In Florida, charter schools initially focused upon at-risk students, but in more recent years, charter schools have diversified to a demographic mix that is similar to the rest of Florida's public schools.

We did not include students who entered and exited multiple schools during 2003-04. Further, because we required at least three consecutive years of data per student, students who entered and exited the Florida public school system during 2001-02, 2002-03, or 2003-04 were also not included.

Separate models for subject and grade level. Learning gains differ by grade level. The typical elementary school student makes larger annual gains in FCAT developmental scale scores than the typical high school student. To account for these differences, we evaluated high schools (students taking ninth- and tenth-grade FCATs), middle schools (students taking sixth-, seventh-, and eighth-grade FCATs), and elementary schools (primarily students taking fifth-grade FCATs and a small percentage of students taking fourth-grade FCATs who had been held back) in separate hierarchical linear models. We produced separate math and reading models for students in each of these three school levels. Exhibit 3 below presents the results of the six models.

Exhibit 3
HLM Model Coefficients for Six Models

	HLM Models					
	High School Math	High School Reading	Middle School Math	Middle School Reading	Elementary School Math	Elementary School Reading
Reliability P0	0.532	0.430	0.533	0.510	0.547	0.453
Reliability P1	0.082	0.081	0.016	0.034	0.107	0.036
Reliability B0	0.707	0.681	0.718	0.697	0.718	0.657
Reliability B10	0.571	0.539	0.724	0.610	0.787	0.602
B00 Coefficient						
G000 Coefficient	1204.688*	1127.917*	966.407*	706.835*	922.522*	855.541*
G001 Coefficient	11.776	21.771	-8.781	8.133	-4.352	-17.282
B01 Coefficient	-65.473*	-112.000*	-107.738*	-95.317*	-80.302*	-128.975*
B02 Coefficient	49.950*	76.085*	85.511*	103.629*	80.668*	84.446*
B03 Coefficient	-12.453*	-45.105*	0.228	-13.675*	-10.135*	-18.944*
B04 Coefficient	-14.746*	-24.985*	-28.219*	-31.121*	-28.794*	-41.562*
B05 Coefficient	-39.184*	-59.628*	-42.728*	-46.219*	-38.009*	-44.507*
B06 Coefficient						
G060 Coefficient	0.557*	0.575*	0.615*	0.645*	0.602*	0.590*
G061 Coefficient	-0.020*	-0.033*	-0.008	-0.012*	-0.011	0.001
B10 Coefficient						
G100 Coefficient	95.214*	20.977*	115.211*	163.788*	177.445*	129.539*
G101 Coefficient	12.440*	28.988*	-1.010	-4.183	-27.411*	-6.894
B11 Coefficient	5.609*	7.261*	3.838*	-0.845	-11.728*	16.482*
B12 Coefficient	2.149*	23.163*	-9.499*	-1.874*	-12.662*	35.107*
B13 Coefficient	3.590*	25.513*	11.222*	16.024*	20.215*	5.100*
B14 Coefficient	-0.589	-10.366*	1.122*	-2.941*	0.975	1.520
B15 Coefficient	0.755	11.518*	8.261*	-0.378	-4.374*	-7.641*
B16 Coefficient						
G160 Coefficient	-0.034*	0.021*	-0.026*	-0.047*	-0.017*	-0.051*
G161 Coefficient	-0.012*	-0.021*	-0.004	0.003	0.003	-0.002
N (level 1)	748,062	783,493	1,201,863	1,295,139	340,976	340,607
N (level 2)	303,326	317,145	466,901	467712	166,236	166,057
N (level 3)	814	812	1,084	1,092	1,865	1,870

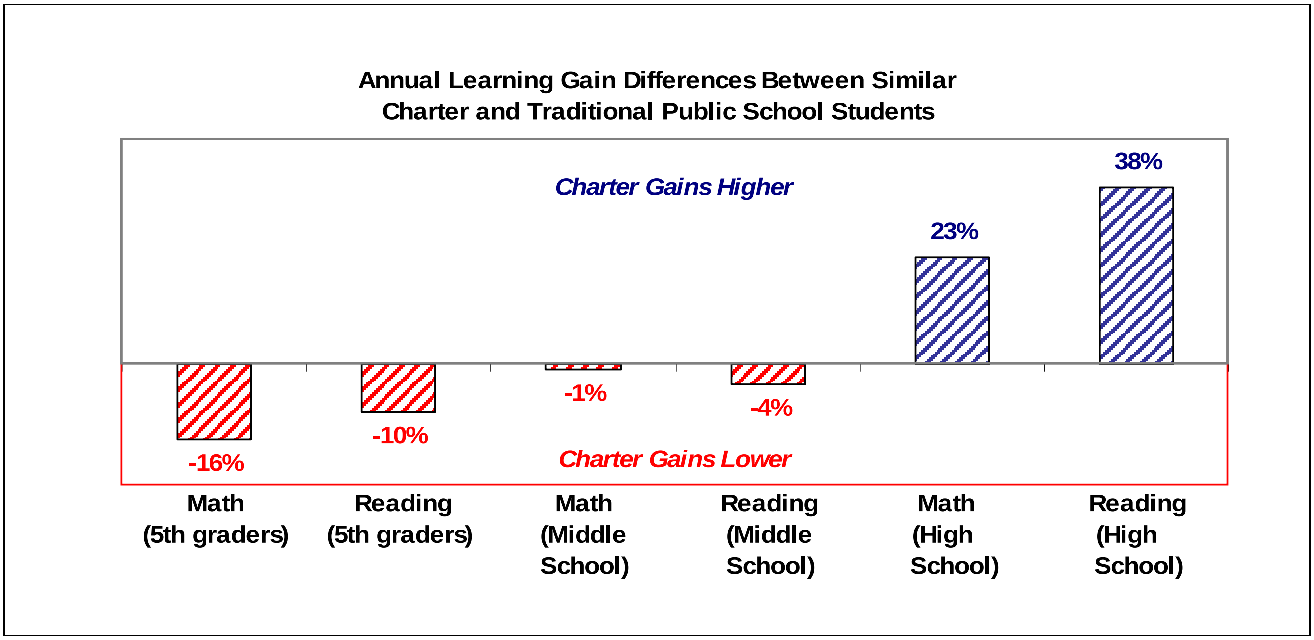
*Significant at the 0.05 level.

Source: OPPAGA analysis of FCAT data, 1998-99 – 2003-04.

Model Interpretation

Charter Effect by Grade Level and Subject. Exhibit 4 below shows the differences between charter and traditional students' FCAT DSS learning gains calculated as a percentage of total average annual learning gains. The differences in learning gains depicted in Exhibit 4 are the coefficients for G101 in the level 3 equations divided by the average annual learning gain of the students in the model. These coefficients are estimates for the difference in learning gains between charter and traditional school students with the same demographic characteristics as a percentage of what the typical student can be expected to gain in a year. A student's total estimated learning gains will vary as a function of the student's baseline FCAT DSS score (see discussion for Exhibits 5 and 6 below).

Exhibit 4
Differences in the Annual Learning Gains of Charter and Traditional Public School Students



Note: Middle school math/reading and fifth-grade reading are not statistically significant at the 0.05 level. See cautions below for discussion of fifth-grade results. Annual learning gain differences between charter and traditional public school students was calculated as a percentage point difference of total average annual learning gain.

Source: OPPAGA analysis of Department of Education data.

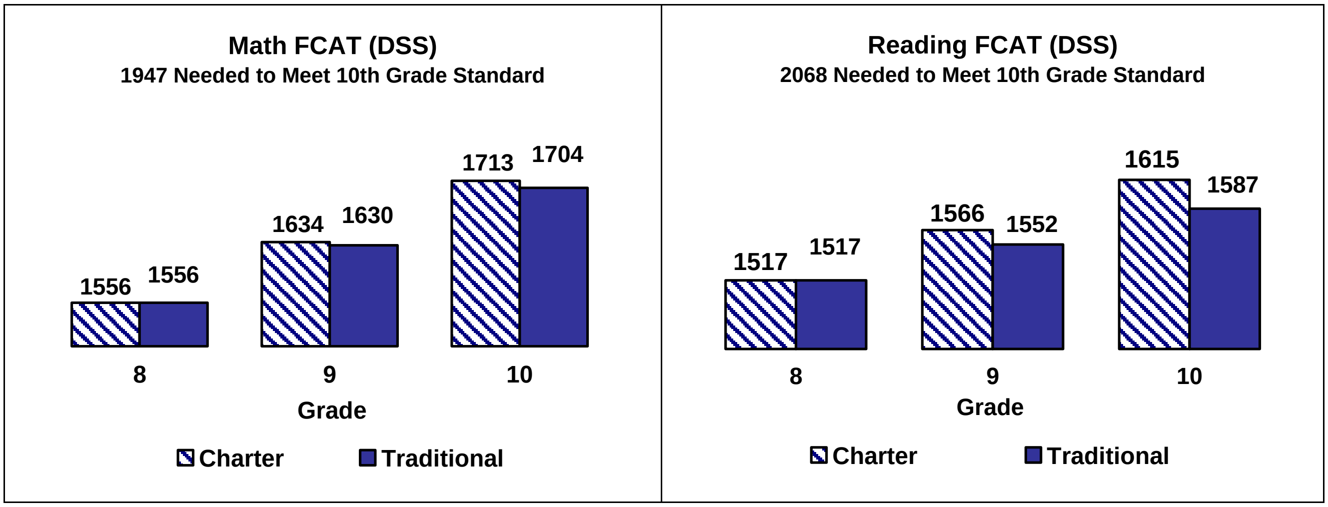
Charter and Traditional Schools' Performance Differences. Exhibit 5 depicts HLM estimates of the predicted learning gains of low-performing eighth graders who stayed in a traditional public school versus entering a charter high school with the same FCAT scores. The developmental scale scores presented in Exhibit 5 are computed using the level 1, level 2, and level 3 equations and the model coefficients in Exhibit 3. As an example, Exhibit 6 shows these calculations for a charter school student's developmental scale score in grade 10 reading.

A student's estimated developmental scale score is calculated using the level 1 equation, Y = P0 + P1*(TIME). "TIME" is the number of years a student is in a school.

The level 2 equations define P0 and P1 as functions of B00 through B06 and B10 through B16 and a centered baseline developmental scale score. Exhibit 5 reflects values for B01-B05 and B11-B15 that are equal to 0. The students' baseline developmental scale scores were centered by subtracting 569 for math and 474 for reading. These centered values represent a baseline developmental scale score of 1200 in math and 1150 in reading, i.e., a high school student who was a low performing seventh grader.

The level 3 equations define B00, B06, B10, and B16 as functions of G000, G001, G060, G061, G100, G101, G160, and G161.

Exhibit 5
Estimated Learning Gains for Low-Performing Students in Charter and Traditional Public High Schools



Source: OPPAGA analysis of Department of Education data.

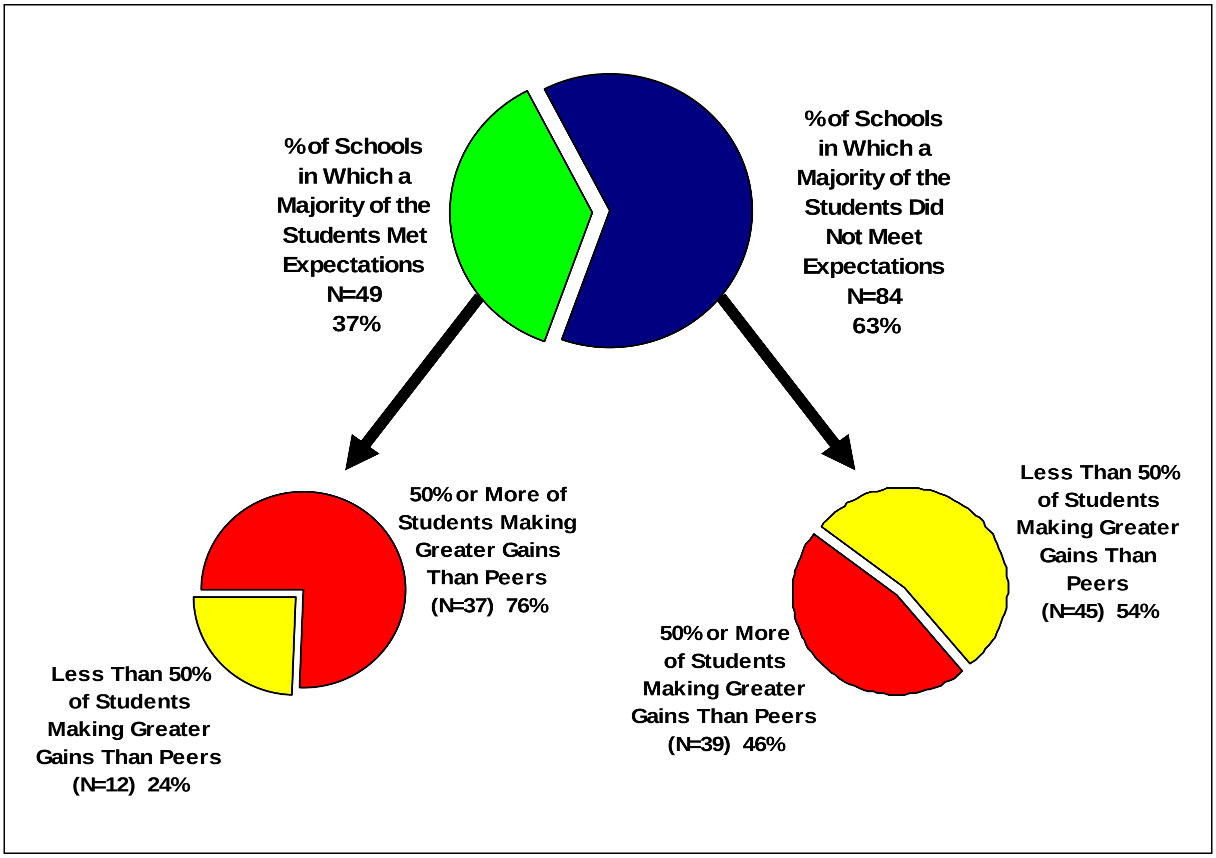
Exhibit 6
Calculation of Reading FCAT Developmental Scale Score for Charter Student in 10th Grade

HLM Equations		Calculation of Reading Developmental Scale Score	
LEVEL 1	Y = P0 + [P1 H (TIME)]	**LEVEL 1**	1615.42271 = 1516.51867 + [49.45202 H (2)]
LEVEL 2	P0 = B00 + [B06 H (FCAT STATUS t-1)]	**LEVEL 2**	1516.51867 = 1149.68727 + [0.54265 H (1150-474)]
	P1 = B10 + [B16 H (FCAT STATUS t-1)]		49.45202 = 49.96578 + [-0.00076 H (1150-474)]
LEVEL 3	B00 = G000 + [G001 H (CHARTER)]	**LEVEL 3**	1149.68727 = 1127.91655 + [21.77072 H (1)]
	B06 = G060 + [G061 H (CHARTER)]		0.54265 = 0.57542 + [-0.03277 H (1)]
	B10 = G100 + [G101 H (CHARTER)]		49.96578 = 20.97746 + [28.98832 H (1)]
	B16 = G160 + [G161 H (CHARTER)]		-0.00076 = 0.02061 + [-0.02137 H (1)]

Source: OPPAGA analysis of Department of Education data.

Describing the performance of charter schools. Exhibit 7 shows that the student performance of charter middle and high schools varies widely. For this analysis we divided charter middle and high schools into two groups—those schools where a majority of their students scored at achievement level 3 on the math or reading FCAT SSS and those schools where a majority of their students scored below achievement level 3 on the math and reading FCAT. We then further divided each group according to whether a majority of a school's students made greater learning gains than similar students statewide. We did not include schools that tested fewer than 25 students. We used a two-step process to calculate the percentage of students that made greater learning gains than similar students statewide. First, we ran the four HLM models for middle and high charter schools without specifying charter status at level 3. We saved the standardized residuals from the level 2 equation for students' estimated annual learning gains (P1). Second, for each student we averaged their math and reading standardized residuals. A positive mean standardized residual indicated that the student was making greater gains than like students statewide.

Exhibit 7
Some Middle and High Charter Schools Are More Successful Than Others at Ensuring Their Students Make Progress Towards State Academic Standards



*Schools with fewer than 25 students tested are not included. N=133.

Source: OPPAGA analysis of Department of Education FCAT data, 1998-99 through 2003-04.

We did a similar analysis of fifth graders in charter elementary schools but did not include the results in Exhibit 7 because fifth graders are a small proportion of elementary school students and may not represent the overall student performance in the school.

Additional statistical analyses, site visits, and caveats

Results from additional statistical analyses supported our HLM findings. In developing the HLM model described above, we conferred with education researchers in the Florida Department of Education and at Florida State University. We tested numerous model specifications prior to deciding on the model presented in this technical report. In all but one model specification that we tested, the charter effect upon high school reading and math was statistically significant. In the one model specification in which we used "grand mean" centered the baseline FCAT DSS score, the differences in learning gains in the high school models were statistically no different from zero, indicating that at the very least high school students in charter schools achieved the same learning gains as students in traditional public high schools (see prior discussion, p. 5, "Alternative specifications of students' baseline developmental scale score").

We conducted an ordinary least squares residual analysis that was independent of our HLM analyses. Our OLS model depicted charter school student learning gains between 2002-03 and 2003-04 as a function of student demographic characteristics and past FCAT performance. We used this analysis to identify charter schools to visit in which the students were making significantly greater or lesser learning gains than similar students in other charter schools. We chose to visit charter schools that had the largest negative and positive residuals, more than plus or minus two standard deviations (i.e., schools with students making the greatest and least learning gains compared to similar charter school students). The high- and low-performing charter schools that we identified in our OLS analysis of charter schools matched the high and low performing charter schools that we identified through our analysis of HLM residuals. The HLM analysis utilized both charter and traditional schools.

Site visits documented different practices among high and low performing charter schools. The results from our statistical models were supported from findings of our sites visits to a sample of 15 charter schools. We conducted site visits to these schools to identify differences in how the higher- and lower-performing schools operated. As described above, we classified charter schools as higher or lower performing based on the annual learning gains of their students. Of the 15 schools, 9 were high performing in terms of producing high student achievement while 6 were low performing. The site visits were conducted on a blind basis as team members were not aware in advance as to whether the schools they visited ranked high or low in performance. The team members rated the schools on criteria outlined in the effective schools literature (Lezotte, 1991).[2] In every case, the teams' ratings of the schools matched the statistical designation of the school as higher or lower performing.

Caveats for math elementary results. Because the hierarchical linear model we used to estimate learning gains required three years of data for each student, our findings for elementary school math learning gains are based primarily on students who were fifth graders during 2003-04. A small percentage of fourth graders who had been held back were also in the analysis. Fifth graders comprise a small proportion of all elementary school students. Our middle school and high school results include students in all grades related to school level (e.g., middle school utilizes grades 6, 7, and 8). We emphasize that the statistically significant negative charter effect for charter school fifth graders in math does not apply to all elementary students. Finally, the size of the negative charter school effect (-27 DSS points) is small relative to the typical learning gains in math that fifth graders achieve (165 DSS points).

Caveats for high school results. Our model best explains the learning gains of those students who have low and middle baseline developmental scores (FCAT SSS achievement levels 1-3) which is the majority of students. Our positive charter findings for high school stem from lower performing charter students outperforming their traditional school peers. We also used separate hierarchical linear models to examine only high performing students (students with an initial baseline score that corresponded to an achievement level of 4 or 5). For students in these levels, there was not a statistically significant difference between the gains of charter and traditional school students.

---

[2] Lezotte (1991). *Correlates of Effective Schools: The First and Second Generation.*

The Florida Legislature

Office of Program Policy Analysis and Government Accountability

Visit the *Florida Monitor*, OPPAGA's online service. See www.oppaga.state.fl.us. This site monitors the performance and accountability of Florida government by making OPPAGA's four primary products available online.

- OPPAGA publications and contracted reviews, such as policy analyses and performance reviews, assess the efficiency and effectiveness of state policies and programs and recommend improvements for Florida government.
- Performance-based program budgeting (PB^2) reports and information offer a variety of tools. Program evaluation and justification reviews assess state programs operating under performance-based program budgeting. Also offered are performance measures information and our assessments of measures.
- Florida Government Accountability Report (FGAR) is an Internet encyclopedia of Florida state government. FGAR offers concise information about state programs, policy issues, and performance.
- Best Financial Management Practices Reviews of Florida school districts. In accordance with the *Sharpening the Pencil Act*, OPPAGA and the Auditor General jointly conduct reviews to determine if a school district is using best financial management practices to help school districts meet the challenge of educating their students in a cost-efficient manner.

Subscribe to OPPAGA's electronic newsletter, Florida Monitor Weekly, a free source for brief e-mail announcements of research reports, conferences, and other resources of interest for Florida's policy research and program evaluation community.

OPPAGA supports the Florida Legislature by providing evaluative research and objective analyses to promote government accountability and the efficient and effective use of public resources. This project was conducted in accordance with applicable evaluation standards. Copies of this report in print or alternate accessible format may be obtained by telephone (850/488-0021 or 800/531-2477), by FAX (850/487-3804), in person, or by mail (OPPAGA Report Production, Claude Pepper Building, Room 312, 111 W. Madison St., Tallahassee, FL 32399-1475). Cover photo by Mark Foley.

Florida Monitor: www.oppaga.state.fl.us

Project conducted by Steve Harkeader, Mark Frederick, and Mark West,
under the supervision of David D. Summers (850/487-9257)

Jane Fletcher, OPPAGA Education Staff Director (850/487-9255)

Gary R. VanLandingham, OPPAGA Director