V. The Longitudinal Analysis

Project STAR researchers hoped that enough students would remain in the study to allow a strong longitudinal analysis. Although each year of the study included more than 6000 students, only 1842 were in the same class-size condition for all four years (K-3; 1985-1989) of the study. Table V-1 shows the data base available for a four year longitudinal analysis. At the end of kindergarten there were no differences between results of students in regular and regular with aide classes, there was parent pressure to reassign some students, and as kindergarten was not mandatory in Tennessee there was a fairly large influx of new students. Students in regular and regular with aide classes were reassigned at random; students in small classes were not reassigned. This reassignment reduced the number of students who met the conditions for the longitudinal analysis, and newly entering students would be excluded as they lacked kindergarten scores. Thus, researchers decided to do a longitudinal analysis that had two parts: Kindergarten-Grade 1 (K-1) and Grades 1, 2, and 3 (1-3). This decision produced more students, schools and classes for the analyses. (See Table V-1)

TABLE V-1

Number of Schools, Students and Classes by Type, Longitudinal Data Base: STAR, 1985-1989*

Groups			Classes							
	Schools	Students	Small		Regular		Regular/Aide		Total	
	N	N	N	%	N	%	N	%	N	%
K-3**	54	1842	91	44	51	25	65	31	207	100
K-1	74	2416	115	38	102	33	90	29	307	100
1-3	60	2571	99	42	64	27	73	31	236	100

*In STAR in the same class type, for 4 years (K-3), or K-1 and 1-3.

**The K-3 analysis tables are in Appendix F. Those tables may be use for gross comparisons. Results are similar, but there are some noticeable differences.

To be considered in the original projected longitudinal analysis, a student had to be in the project all four years, starting in kindergarten (K), be in the same class type (small, regular or regular with a full-time teacher aide) for the entire project, and have the appropriate test scores needed for the analysis. The revised analyses (K-1, 1-3) held to the same general rules: a student was in the study for the requisite number of years and had to have all of the required test data points.

Some Caveats

The original sample (more than 6,000 students in approximately 100 classes of each of the three types) was drawn within the limits of funding and with hopes that there would remain in the study, in their same class types, enough students for a definitive longitudinal result. Throughout the study, as students moved they were replaced by other students placed at random into the three class conditions. From the original substantial data base, only 1842 students of the original kindergarten sample of 6328, or 29 percent, met all conditions for the longitudinal analysis. Using K-1 and 1-3 as longitudinal analyses bases provided more students (2416 and 2571 respectively), but even with the increase in Grade 1 the 1-3 group was still only about 34 percent and the K-1 was 38 percent.

Results of the longitudinal analysis presented here should be treated as tentative due to the restricted subsamples (about one-third of the total group) in each analysis. These youngsters may not be typical of the entire project population. Each subsample of students was divided into classes to obtain class averages for analyses. The absolute number of students for the longitudinal analysis in each of the three conditions was Small (K-1: 1140; 1-3: 891), Regular (K-1: 663; 1-3: 744), and Regular/Aide (K-1: 613; 1-3: 936). Some "large" classes had only 3, 4, 5 or 6 students who constituted the average for that class for the analysis (the same is true for other class types).

Tests of Significance of Mean Difference

Each longitudinal analysis was done for 4 measures -- Word Study Skills (WSS), Total Reading (Read), Total Mathematics (Math) and Total Listening (Listen), and for appropriate testing time points in Kindergarten, Grade 1, Grade 2, and Grade 3.

The total file analysis was a LOCATION X TYPE design, with schools nested in locations and crossed with class types. Grade differences are the dependent variables for multivariate tests of the grade effect (the specific contrasts are K-1 and 1-2, 2-3) and for interactions of grade with location and class type. This design is in Table V-2 for the K-1 analysis. Also, a race file analysis was done, using a LOCATION x RACE x TYPE design with each school having only minority or white students. This was necessitated because there were insufficient degrees of freedom for Schools x Race or Schools x Race x Type in the longitudinal data sets, making it difficult or impossible to test some effects in the completely crossed design. In this analysis, schools with location/race combinations and schools x type are the error terms for every effect of interest. This design is in Table V-3 for the 1-3 analysis.

In this analysis, the class was the unit of measure. In some class-type situations only a few students were in a class for all appropriate years. That is, throughout the project some students moved in and out of the class; only those who remained for the years of the analyses (K-1, 1-3) and had the needed test scores were used to develop the class average used in the analysis. (Appendix F contains the tables showing the numbers for the K-3 analysis with the 1842 students.)

TABLE V-2

Design for Total Class Analysis, Showing the Source of Variation, Error Terms and Degrees of Freedom, Longitudinal Study: STAR 1986-1989, Grades K-1

Source of Variation	Error Term
Grade (G)	Schools by Location (S:L)
LOCATION x GRADE (LG)	S:L
TYPE (T)	S:L
TYPE x GRADE (TG)	TxS:L
LOCATION x TYPE x GRADE (LTG)	TxS:L

	Degrees of Freedom			
	WSS	Reading	Math	Listening
Schools:Location (S:L)	56	56	56	56
Type x Schools (TxS:L)	99	99	100	99

TABLE V-3

Design for Analysis by Race, Showing source of Variation, Error Terms and Degrees of Freedom, Longitudinal Study: STAR 1985-1989

Source of Variation	Error Term
GRADE (G)	SCHOOLSxRACExLOCATION (S:R:L)
LOCATION x GRADE (LG)	S:R:L
TYPE x GRADE (TG)	TxS:R:L
RACE (R)	S:R:L
RACE x GRADE (RG)	S:R:L
LOCATION x RACE x GRADE (LRG)	S:R:L
LOCATION x TYPE x GRADE (LTG)	TxS:R:L
RACE x TYPE x GRADE (RTG)	TxS:R:L
RACE x TYPE (RT)	TxS:R:L
LOCATION x RACE x TYPE x GRADE (LRTG)	TxS:R:L

	Degrees of Freedom			
	WSS	Reading	Math	Listening
Schools:Race:Location (S:R:L)	60	63	63	63
Type x Schools:Race: Location (TxS:R:L)	93	98	98	98

TABLE V-4

Analysis of Variance Results Expressed as Significance Levels, Project STAR, Longitudinal Analysis (1985-1989) Showing the Total Class Results and the Class Results by Race

	Word Study Skills		Total Reading		Total Math		Total Listening		
	K-1	1-3	K-1	1-3	K-1	1-3	K-1	1-3	
GRADE	.01	.001	.01	.001	.01	.001	.01	.001	[A]*
LOC X GRADE	.01	N.S.	.01	.01	N.S.	N.S.	---	N.S.	[B]
TYPE	.01	.001	.01	.001	.01	.001	---	.01	[C]
TYPE X GRADE	.05	N.S.	.01	N.S.	N.S.	N.S.	---	N.S.	[D]
LOC X TYPE X GRADE	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	---	N.S.	[E]
RACE	---	.01	---	.001	---	.01	---	.01	[F]
RACE X GRADE	.01	N.S.	.05	N.S.	N.S.	N.S.	---	N.S.	[G]
RACE X LOC X GRADE	.05	N.S.	N.S.	N.S.	N.S.	N.S.	---	N.S.	[G]
RACE X TYPE	---	N.S.	---	N.S.		N.S.	---	N.S.	[H]
RACE X TYPE X GRADE	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	---	N.S.	[I]
RACE X LOC X TYPE X GRADE	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	---	N.S..	[I]

*Results are discussed on the following pages using the designations [A]-[I] to identify the results being discussed.

N.S.=Not Significant; significance levels p<.001, .01, or .05 reported.

Discussion of the Longitudinal ANOVA Results (Tables V-4 and V-5)

[A] There was statistically significant student growth on the standardized tests on all four measures and at all grade levels. This does not address class size.

[B] There were no statistically significant differences in student growth between/among the classes in the various locations (Inner City, Suburban, Rural, Urban) except in total reading for the 1-2-3 analysis where inner-city gained significantly more from G1 to G2 and from G2-G3 than all other locations. In the K-1 analysis, there were statistically significant gains between/among class types in locations, with inner-city students gaining most in Total Reading and Word Study Skills. This result does not address class size and is shown in Table V-5. Note also that Table V-5 shows that the gain in all locations was fairly similar, with a range of 77.2 to 105.7 favoring the inner city. Inner city class results were consistently the lowest and, except for K, rural classes had the highest results. Note also (Table V-5) that the largest difference between inner city score (lowest) and the highest score in any given year fluctuates from 49.4 to 24.4 with the superior gains in the inner city in G1-G2 and G2-G3 reducing the differences.

[C] Small-regular contrast is significant on all scales at or beyond $p<.01$; aide-regular contrast is not significant for any scale.

[D] There was no interaction with class type over years 1-3 of the study. All class types grow equally, on the average. That is, the small-class advantage which originated in K neither increased or decreased in a statistically significant manner over the subsequent three years.

[E] There were no statistically significant Location x Class Type x Grade interactions on any measures.

[F] Race effects (1-3) significant on all scales at or beyond $p<.01$. Whites do better than minorities on all these measures. K-1 analysis was not run.

[G] In general, grade-to-grade growth in 1-3 was similar for whites (W) and minorities (M), although the differences for the average scores for W and M were considerably less on all four measures for small classes than for the other two class types. In K-1, whites' gains, on average, exceeded gains for minorities on word study and reading. Generally, grade-to-grade growth was the same for whites and minorities, regardless of location.

[H] No statistically significant differential impact of small classes on whites or minorities.

[I] There is no evidence of a differential impact of small classes on whites or minorities, as small classes affect "growth" in each year equally. That is, there may be differential impact on end-of-year performance (see Chapter IV) but not on the total amount of change from K to 1, or 1 to 2 to 3 when students in the project are considered over time. There is no significant Race x Location x Class Type x Grade interation. However, since there were only a few locations (i.e., school types) that had both white and minority students, the test of this effect is based on very small segments of the data.

TABLE V-5

Total Reading Mean Scores by Location
STAR, 1985-1989

	K-1 Analysis			Grades 1-2-3 Analysis					
	K	Gain K-G1	G1	G1	Gain G1-G2	G2	Gain G2-G3	G3	Gain G1-G3
Inner-city	433.3	58.3	491.6	496.8	67.6	564.4	38.1	602.5	105.7
Suburban	468.1	63.6	531.7	535.8	57.1	592.9	27.5	620.4	84.6
Rural	440.9	94.4	535.3	546.2	56.8	603.0	23.9	626.9	80.7
Urban	447.3	89.3	536.6	542.5	53.5	596.0	23.7	619.7	77.2
	34.8*			38.6*		49.4*		24.4*	

*Largest difference between Inner City and any other group.

Longitudinal Average Scores By Grades, By Tests, By Class Types

Tables V-6 through V-8 show the average scores and totals for the three class types by the four locations and for grade levels K-1 and 1-3 for the two measures available for longitudinal analysis: total reading and total math. These tables also show the average growth: K-1, 1-2, 2-3 and the total growth 1-3 for each class type. Except for the scores in brackets in each table, the small class average score exceeds the average scores of other class conditions. Figures V-1 and V-2 show the average annual SAT scaled scores by class type.

Longitudinal Average Scores by Race by Class Type and Total

Table V-8 shows that on all four measures the differences between average scores of Minorities (M) and Whites (W) are far less in small classes than in regular and regular/aide classes. Average scores for (M) are considerably higher in small classes than for (M) in the other two class types, and although the average scores for (W) in small classes are higher than average scores of (W) in the other two class types, the differences are not as extreme as for (M). Minority students in small classes outperform minority students in other class types and very nearly equal the performance of white students in regular and regular/aide classes.

TABLE V-6

Average Annual Scores, Year-to-Year Growth by Class Type and Location, Longitudinal Analysis, Project STAR (1985-1989): Total Reading

	SMALL					REGULAR					REGULAR/AIDE				
	K-1 Analysis		1-2-3 Analysis			K-1 Analysis		1-2-3 Analysis			K-1 Analysis		1-2-3 Analysis		
	K	G1	G1	G2	G3	K	G1	G1	G2	G3	K	G1	G1	G2	G3
Inner City	436.4	502.9	506.1	576.6	610.0	431.1	481.5	482.6	543.5	589.4	431.8	488.3	494.5	563.4	602.2
Suburban	448.8	545.4	545.5	599.0	626.7	436.3	518.2	531.3	592.8	617.3	435.4	527.7	526.0	583.8	614.2
Rural	444.9	542.8	548.8	606.5	631.8	439.2	526.1	540.4	598.2	623.1	437.8	536.0	547.7	602.7	624.2
Urban	451.3	545.7	550.1	598.7	623.7	444.3	530.7	542.4	595.7	619.6	445.4	531.6	532.9	592.7	614.5
All	444.8	535.7	540.4	598.5	625.9	437.6	515.8	529.1	588.0	616.0	437.3	525.3	533.5	591.7	617.7

	K-1	1-2	2-3	K-1	1-2	2-3	K-1	1-2	2-3
Growth/Year	90.9	58.2	27.3	78.1	58.9	28.0	88.0	58.2	26.0

	SMALL	REGULAR	REGULAR/AIDE
Total Growth (1-3)	85.5	86.9	84.2

TABLE V-7

Average Annual Scores, Year-to-Year Growth by Class Type and Location, Longitudinal Analysis, Project STAR (1985-1989): Total Math

	SMALL					REGULAR					REGULAR/AIDE				
	K-1 Analysis		1-2-3 Analysis			K-1 Analysis		1-2-3 Analysis			K-1 Analysis		1-2-3 Analysis		
	K	G1	G1	G2	G3	K	G1	G1	G2	G3	K	G1	G1	G2	G3
Inner City	484.9	502.9	525.6	574.4	612.7	477.6	[503.9]	505.1	547.0	592.1	472.0	[508.0]	512.2	562.7	604.0
Suburban	504.2	550.0	552.3	598.3	627.1	495.9	526.8	538.6	582.7	619.8	488.2	531.2	530.8	576.3	611.1
Rural	499.4	543.6	552.1	597.6	632.1	487.0	535.6	545.5	593.9	627.2	489.7	541.1	548.3	[597.8]	628.0
Urban	500.0	547.5	548.9	585.3	617.2	487.8	536.4	[550.0]	585.2	622.6	496.7	535.4	528.6	582.5	613.4
All	497.5	541.8	547.0	592.4	626.0	486.5	527.2	537.7	582.9	619.4	487.1	532.7	537.1	586.4	619.3

	K-1	1-2	2-3	K-1	1-2	2-3	K-1	1-2	2-3
Growth/Year	44.3	45.4	33.6	40.7	45.2	36.4	45.6	49.2	32.9

	SMALL	REGULAR	REGULAR/AIDE
Total Growth (1-3)	79.0	81.6	82.1

TABLE V-8

Average Annual Scores and Differences Between the Scores of White (WH) and Minority (MIN) Students By Class Type and Total on Two Measures Longitudinal Analyses: Project STAR, 1985-1989, K-1 and 1-2-3

TOTAL READING

	SMALL REGULAR				REGULAR/AIDE				TOTAL			
	K-1 Analysis		1-2-3 Analysis		K-1 Analysis		1-2-3 Analysis		K-1 Analysis		1-2-3 Analysis	
	K-1 Analysis		1-2-3 Analysis									
	K	G1	G1	G2	G3	K	G1	G1	G2	G3	K	G1
	G1	G2	G3	K	G1	G1	G2	G3				
WH	449.8	542.4	609.2	631.5	444.9	537.0	604.3	625.8	444.0	536.9	606.4	626.2
	446.6	539.1	607.0	628.2								
MIN	438.0	517.0	582.2	612.7	419.4	504.9	559.2	596.0	423.8	498.4	569.1	596.9
	431.2	509.9	575.0	605.8								
DIF	11.8	25.4	27.0	18.8	25.5	32.1	45.1	29.8	20.2	38.5	37.3	29.3
	15.4	29.2	32.0	22.4								

TOTAL MATH

	SMALL REGULAR				REGULAR/AIDE				TOTAL			
	K-1 Analysis		1-2-3 Analysis		K-1 Analysis		1-2-3 Analysis		K-1 Analysis		1-2-3 Analysis	
	K-1 Analysis		1-2-3 Analysis									
	K	G1	G1	G2	G3	K	G1	G1	G2	G3	K	G1
	G1	G2	G3	K	G1	G1	G2	G3				
WH	502.9	548.2	555.0	600.2	632.4	489.4	538.5	546.9	594.0	626.6	492.7	541.0
	545.6	594.6	626.2	495.6	543.0	549.8	596.7	628.8				
MIN	485.2	524.2	530.9	577.3	613.0	481.5	507.0	517.6	554.9	597.2	476.0	514.4
	520.5	567.9	602.3	481.7	515.8	523.8	567.8	605.1				
DIF	17.7	24.0	24.1	22.9	19.4	7.9	31.5	29.3	39.1	29.4	16.7	26.6
	25.1	26.7	23.9	13.9	27.2	26.0	28.9	23.7				

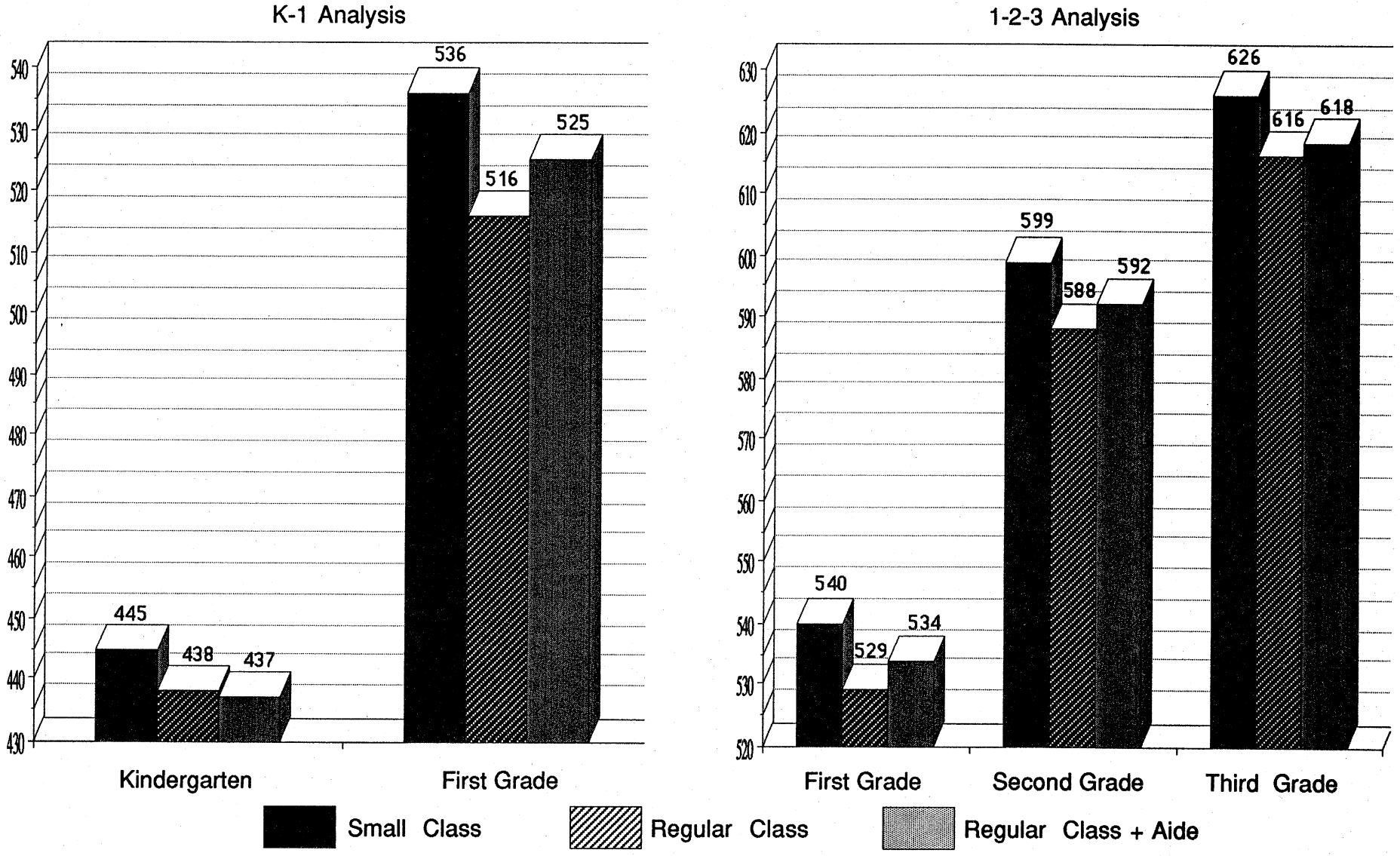
WH=WHITE MIN=MINORITY DIF=DIFFERENCE

Figure V-1

Project STAR

Average Annual Scaled Scores, Year-to Year Growth by Class Type

Longitudinal Analysis, (1985-1989): Total Reading



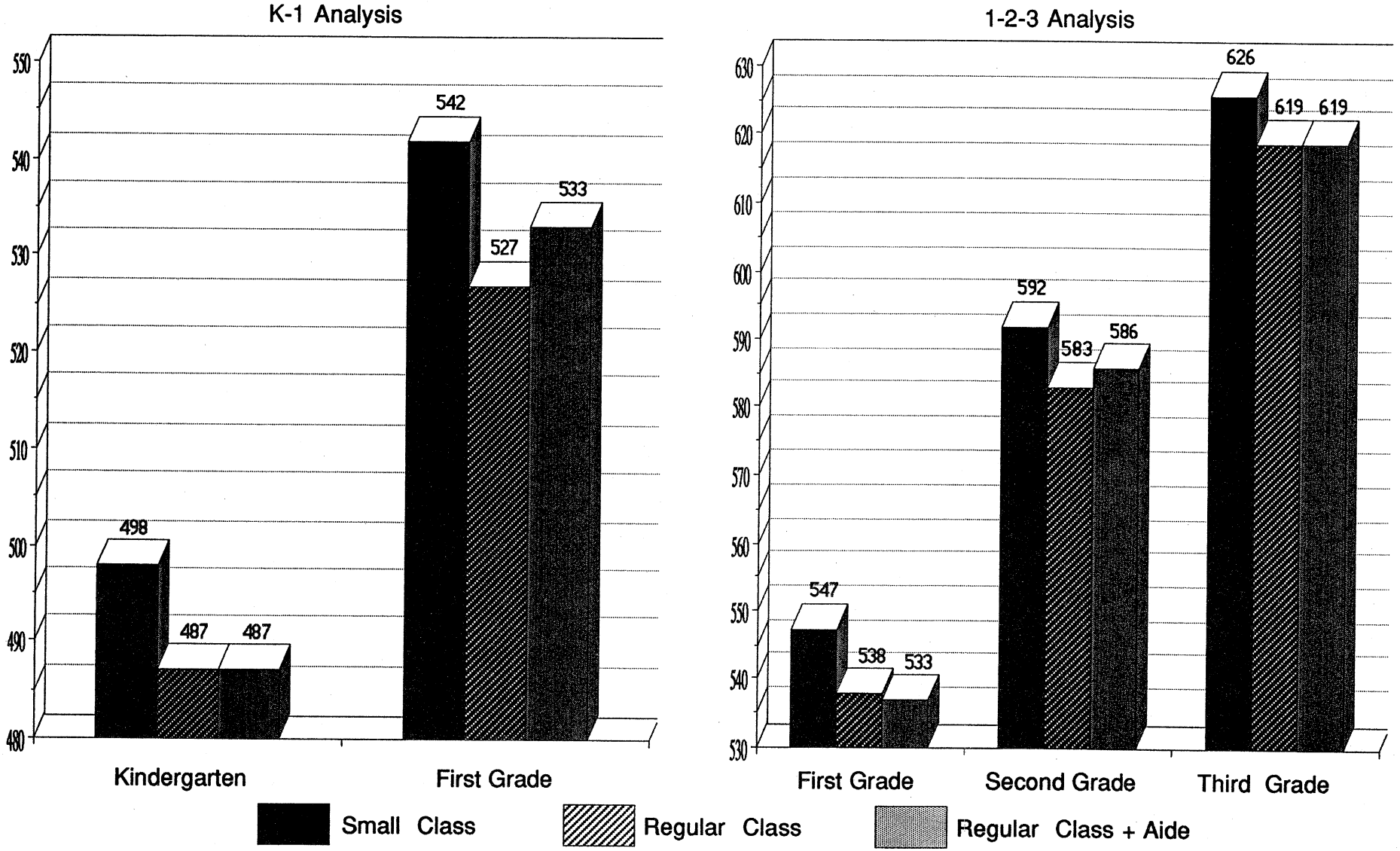
Stanford SESAT II, Primary I, II, and III

Figure V-2

Project STAR

Average Annual Scaled Scores, Year-to Year Growth by Class Type

Longitudinal Analysis, (1985-1989): Total Math



Stanford SESAT II, Primary I, II, and III

SUMMARY COMMENTS

Although each yearly analysis continued to identify the benefits of a student's being in a small class, the results for the small (about 33 percent) subsample of students in the same class size for 2 years (K-1) and 3 years (1-3) were less definitive for student achievement. The results showed that the large and statistically significant gains favoring the small classes made in the first year (i.e., K in the K-1 comparison and Grade 1 in the 1-3 comparison) were maintained, but that there were no statistically significant gains in future years. Likewise, the average scores on the four measures of achievement (detailed tables provided for Total Reading and Total Math only) used for the longitudinal analyses showed that the minority students in small classes achieved higher scores than minority students in the other class conditions, but the non-minority students continued to outperform the minority students in all class types and locations.