II. STUDY DESIGN

A. Introduction

The STAR design had to provide answers for questions required by the legislation, meet requirements for a longitudinal study, review one-year or cross-sectional effects of the treatment, and answer questions of interest. Two key design decisions were to have a within-school design and random assignment of both teachers and students to class types. STAR was a randomized experiment conducted **in situ**. The control-group design was Campbell and Stanley (1963) Design Number 6, a randomized experiment employing post-test analysis only. The primary analysis was built on post-test only design. Additional analyses employed other analytic models.

B. Choice of Within-School Design

Because of potentially large differences between schools (i.e., school effects) in such items as resources, teachers and students, the consortium chose a within-school design. A within-school design reduced major sources of possible variation in student achievement attributable to school effects. This decision required that each school have sufficient enrollment in each grade (at least 57 students) to provide at least one small (13-17 enrollment), one regular (22-25), and one regular with a full-time aide (22-25) class. In schools with larger enrollments, additional classes were established. This design assured that there would be the same kinds of students, curriculum, principal, policy, schedule, expenditures, etc., for each class type by school and avoided the problem of control groups that were not motivated to attend carefully to project needs since they probably would gain nothing by remaining in the project. In the within-school design the control classes participated fully in all testing, etc., since it was part of the project. An entire project school might do better than expected due to project participation (the halo or Hawthorne effect). Reciprocally, it was also possible that competition could occur within the school whereby the control teacher(s) would work extra hard (the John Henry effect).

After initial selection of participating systems, the choice of schools within systems was partly a function of school size. Grade-level enrollment determined the number of classes of each type established in each school. For example, the 79 schools selected to participate in Project STAR (kindergarten) provided enough classes (small, regular, regular w/aide) to meet the design estimate of approximately 100 classes of each type.

C. Selection into the Three Conditions

The 79 project elementary schools selected in the first year served rural, urban, suburban and inner-city students. The within-school design required each participating school to have three or more classes. Larger schools had more classes distributed among the three class types. Table II-1 shows the design configurations for establishing classes in schools of various sizes. A student in a small class in kindergarten remained in the small class for grades one, two and three, to assist the measurement of cumulative effect of the class type. In kindergarten (1985-1986), there were 128 small classes, 101 regular classes, and 99 regular classes with full-time teacher aides. Approximately 6,500 students participated in Project STAR in kindergarten.

TABLE II-1

Plan for Distribution of Students and Classes in Within-School Design: Project STAR (1985-1986)

Design Type	Enrollment (ADM)	Classes (N)	Class Types	Extra Room Needed
One	57-67	(3)	S,R,R/A	No
Two	68-78	(4)	S,S,R,R/A	Yes
Three	79-92	(4)	S,R,R/A,R/A or S,R,R,R/A	No
Four	93-109	(5)	S,S,R,R,R/A or S,S,R,R/A,R/A	Yes
Five	110-134	(6)	S,S,R,R,R/A,R/A	Yes
Six	135+	(7+)	Individually Designed	Yes

S=Small Class (1:13-17);R=Regular Class (1:22-25);
R/A=Regular Class with a Full-time Teacher Aide (1:22-25)

The plan described in Table II-1 was used to govern the selection of class condition throughout the study. Once assigned to a class type a student was to remain in the assigned class type as long as he/she was in the project. Due primarily to teacher-identified discipline problems and some parent complaints, the STAR consortium had to revise this procedure after the kindergarten year. Since there were no differences on any measure for students in regular and regular with aide classes, students who had been in these class types in kindergarten were reassigned randomly within the two class types for first grade. The external advisory committee informed STAR that this interchanging could create problems in conducting longitudinal analysis. Therefore, first grade was the only grade in which students in regular and regular with aide classes were permitted to interchange. No further changes were made after first grade. Table II-2 lists STAR schools and systems and shows the location designation and class type design for kindergarten through third grade. Figure II-1 shows how the participating schools were distributed across the state. Table II-3 shows the number of schools and students by location for each year of the study.

D. Modifications in Study Design

In a large-scale field project some changes occur that cannot be anticipated. Schools may drop out of the project; classes will gain or lose students; and in some cases these changes will make a class too small or too large for the design. The researchers took these possibilities into consideration by over designing the project. A power test at the beginning of the project indicated that it would be possible to detect a small achievement difference (3% or more) with only 80 classes of each type, or a total of 240 classes, rather than the 329 that actually participated. At the end of kindergarten, 34 classes had either too many or too few students for

the original design (e.g., a small class may have ended up with 12 students rather than staying within the 13-17 range). Data were analyzed both including and excluding the 34 classes and results of both analyses were substantially the same. Oversampling was necessary because of the expected attrition of students and schools over the project's four years.

TABLE II-2

Project STAR Systems, Schools, and Designs Kindergarten through Third Grade (1985-1989)

System	School	Design			
		Kindergarten	1st Grade	2nd Grade	3rd Grade
		S R A	S R A	S R A	S R A
Bedford	Thomas Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Bledsoe	Pikeville Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Blount	Bungalow Elem.	1-1-1	Withdrawn	Withdrawn	Withdrawn
Blount	Midsettlements	1-1-1	2-1-1	2-1-1	2-1-1
Blount	Rockford Elem	1-1-1	2-1-1	2-1-1	2-1-1
Chester	East Chester Elem.	2-2-1	2-2-2	2-2-2	3-1-2
Claiborne	Ellen Myers Elem.	2-1-1	1-1-2	1-1-2	1-1-2
Claiborne	Tazewell (TNT) Elem.	2-2-2	2-2-2	2-1-2	2-1-2
Clay	Celina Elem.	2-1-1	2-1-1	2-1-1	2-1-1
Coffee	North Coffee Elem.	1-1-1	1-1-1	2-1-1	2-1-1
Cumberland	Homestead Elem.	1-1-1	1-1-1	2-1-1	2-1-1
Cumberland	Crossville Elem.	2-1-2	2-1-2	2-1-2	2-1-2
Davidson	Rosebank Elem.	1-1-1	2-1-1	2-1-1	2-1-1
Davidson	Hattie Cotton Elem.	1-1-1	2-2-1	2-1-1	2-1-1
Davidson	Cole Elem.	3-2-2	3-5-4	3-4-4	3-2-4
Davidson	Andrew Jackson Elem.	1-1-1	2-2-1	3-1-1	3-1-1
Decatur	Parsons Elem.	2-1-1	1-1-2	1-1-2	2-0-2
Dyer	Newbern Elem.	1-1-1	1-2-1	1-2-1	1-2-1
Fentress	York Elem.	2-1-1	2-1-1	2-1-1	2-1-1
Hamilton	Daisy Elem.	2-1-1	1-1-2	1-1-2	1-1-2
Hamilton	Ganns-Mid Valley	2-1-1	1-2-1	1-2-1	1-2-1
Hamilton	Soddy Elem.	1-1-1	2-1-1	2-1-1	2-1-1
Hancock	Hancock Central Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Hardin	Parris South Elem.	2-1-1	2-2-1	2-2-1	3-1-1
Hardin	Savannah North Elem.	1-1-1	1-1-1	2-1-1	3-0-1
Humphreys	Waverly Elem.	2-2-2	3-2-2	3-2-2	3-2-2
Jefferson	Jefferson Elem.	2-2-1	2-2-1	2-2-1	2-2-1
Jefferson	White Pine Elem.	1-1-1	1-1-1	2-1-1	2-1-1
Knoxville	Alice Bell Elem.	1-1-1	2-1-1	2-1-1	2-1-1
Knoxville	Bearden Elem.	2-1-1	Withdrawn	Withdrawn	Withdrawn
Knoxville	Rocky Hill Elem.	2-1-1	1-1-1	2-1-1	2-1-1
Knoxville	Sara Moore Greene	2-1-1	1-1-1	2-1-2	2-1-2
Knoxville	Green Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Lawrence	South Lawrence Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Lawrence	Lawrenceburg Elem.	2-1-1	2-1-1	2-1-1	2-1-1

TABLE II-2 Cont.

System	School	Kindergarten	1st Grade	2nd Grade	3rd Grade
		SRA	SRA	SRA	SRA
Lenoir	Lenoir City Elem.	2-1-2	2-1-2	2-1-2	2-1-2
Lewis	Lewis County Elem.	2-2-2	2-2-2	2-2-2	3-2-2
Macon	Enon Kindergarten	2-2-1	Kindergarten	Only	
Macon	Fairlane Elem. 1-3		2-2-2	2-2-2	2-2-2
Manchester	College Street Elem.	1-1-1	1-1-1	2-0-1	2-0-1
Marion	South Pittsburg Elem.	2-1-1	2-1-1	2-1-1	2-1-1
Maryville	John Sevier Elem.	2-1-1	1-1-2	1-1-2	1-1-2
Maury	Mt. Pleasant Elem.	2-2-1	2-2-2	2-1-2	2-1-2
Maury	Spring Hill Elem.	2-1-1	2-1-1	2-1-1	2-1-1
McNairy	Selmer Elem.	2-2-2	2-2-2	2-2-2	2-2-2
Memphis	Caldwell Elem.	2-1-2	2-2-2	2-2-2	2-2-2
Memphis	Cummings Elem.	2-1-2	2-2-2	2-2-2	1-1-1
Memphis	Double Tree Elem.	1-1-2	2-2-2	2-1-2	2-1-2
Memphis	Douglas Elem. Sch.	1-1-1	2-1-1	2-1-1	1-1-1
Memphis	Florida Elem. Sch.	1-1-1	1-2-1	1-1-1	1-1-1
Memphis	Goodlett Elem. Sch.	1-1-2	2-1-2	1-1-2	1-1-2
Memphis	Gordon Elem. Sch.	1-2-1	1-2-1	1-2-1	1-2-1
Memphis	Hanley Elem. Sch.	2-3-2	3-2-2	3-2-2	3-2-2
Memphis	A. B. Hill Elem. Sch.	2-1-2	2-2-2	2-2-2	2-1-2
Memphis	Kansas Elem. Sch.	1-1-1	1-1-1	1-1-1	2-0-1
Memphis	Larose Elem. Sch.	2-1-2	2-2-2	3-2-2	2-2-2
Memphis	Lester Demo. Sch.	4-3-2	2-2-2	2-1-2	2-1-2
Memphis	Lincoln Elem. Sch.	1-1-1	1-1-1	1-1-1	1-1-1
Memphis	Orleans Elem. Sch.	2-1-1	2-1-1	2-1-1	2-0-1
Memphis	Raineshaven Elem.	2-1-1	2-1-2	2-1-2	2-1-2
Memphis	Raleigh-Bartlett	2-1-1	1-2-1	2-1-1	2-1-1
Memphis	Riverview Elem.	2-3-2	2-3-2	2-2-3	2-2-2
Memphis	Snowden Elem.	2-1-1	Withdrawn	Withdrawn	Withdrawn
Memphis	Westside Elem.	1-1-1	2-1-1	2-1-1	2-1-1
Memphis	Whitehaven Elem.	2-2-1	2-2-1	1-3-1	1-3-1
Montgomery	Montgomery Central	1-2-1	2-1-1	2-1-1	2-1-1
Obion	South Fulton Elem.	2-1-1	1-1-2	1-1-2	1-1-2
Perry	Linden Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Pickett	Pickett Co. Elem.	1-1-1	1-1-1	1-2-1	1-2-1
Rhea	Spring City Elem.	2-1-2	2-2-2	2-1-2	2-1-2
Trenton	Trenton Elem.	2-1-2	2-2-2	Withdrawn	Withdrawn
Trousdale	Trousdale Co. Elem.	1-2-1	1-2-1	2-2-1	3-1-1
Tullahoma	East Lincoln Elem.	2-1-1	1-1-1	2-0-1	2-0-1
Unicoi	Unicoi Elem.	1-1-1	1-1-1	1-1-1	1-1-1
Washington	Jonesborough Elem.	3-2-2	3-2-2	3-2-2	3-2-2
Washington	Boones Creek Elem.	2-2-1	2-2-1	2-2-1	2-2-1
Wayne	Collinwood Elem.	1-1-1	2-1-1	2-1-1	2-1-1
White	Findlay Elem.	1-2-1	2-2-1	2-1-1	2-1-1
Williamson	W. P. Scales Elem.	2-1-1	2-1-1	2-2-1	2-2-1
Wilson	Lakeview Elem.	2-1-2	2-1-2	2-2-2	2-2-2

Figure II - 1

Distribution of Schools Participating in

Project STAR

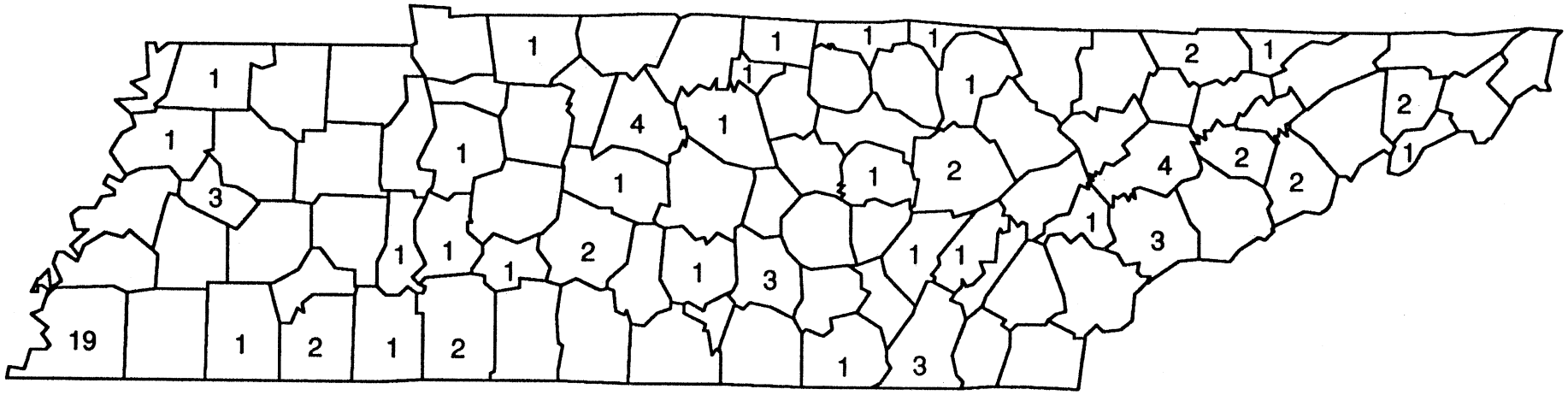

TABLE II-3

Number of Schools and Students by Location Kindergarten through Third Grade (1985-1989)

	Kindergarten (1985-86)		Grade 1 (1986-87)		Grade 2 (1987-88)		Grade 3 (1988-89)	
Location	Students	Schools	Students	Schools	Students	Schools	Students	Schools
Rural	2918	39	3240	38	3168	38	3239	38
Urban	568	8	686	8	482	7	506	7
Suburban	1414	16	1589	15	1711	15	1722	15
Inner City	1428	17	1380	15	1485	15	1336	15
Total	**6328**	**79**	**6835**	**76**	**6846**	**75**	**6804**	**75**

E. Data Collection Instruments

Project personnel collected information (data) about student achievement and development and about variables other than class size that might affect achievement. This included collecting information about instructional processes to try to understand how reduced class size effects and reduced student-adult ratio effects occurred. These effects were examined over time in relation to students, teachers and teacher aides. The impact of reduced class size and reduced student-adult ratio was assessed through multiple measures of student achievement and development and process measures such as activity logs, classroom observation, etc.

1. Tests

a. Stanford Achievement Test (SAT)

Students were tested each spring on the dates specified by the state. In each grade, the appropriate level of SAT was administered to all project STAR students and to students in 21 comparison schools. The norm-referenced SATs cover reading, math, spelling, listening, and in the higher grades science and social science, and provide subscores for both reading and math (The Psychological Corporation, Harcourt Brace Jovanovich, Inc., 1985).

b. Tennessee's Basic Skills First Test (BSF)

The State developed Basic Skills Criterion Tests for the third, sixth and eighth grades in reading and math in 1984. Because the SAT does not cover all of the curriculum taught, and the curriculum does not cover everything that is tested, Project STAR contracted with the State Testing Service to develop STAR Criterion Tests in reading and math to cover BSF learning objectives in grades one and two. These tests were similar to the already developed third grade test. The BSF learning objectives were the criteria tested. The untimed tests consist of multiple choice items, four items per objective, and are designed so that they can be administered in about an hour. (Tennessee Department of Education, 1987).

c. Self-Concept and Motivation Inventory

In addition to the SAT and BSF tests, students completed a self-concept and motivation inventory, the SCAMIN (Person-O-Metrics, Inc., 1967,1968). The SCAMIN asks students to indicate pictorially their response to 24 situations. For example what "face" (i.e. happy, sad, indifferent, etc.) would the student wear if he/she "had to tell his/her parents they lost their coat." The SCAMIN was selected because it is group administered, has forms appropriate for grades K-3, measures elements of self-concept of concern to this project, and requires no special training for administration. While it has only moderate reliability for the early grades, the SCAMIN is useful for comparing groups, such as small classes with regular classes (Davis and Johnston, et al.).

2. Additional Data Collection Instruments

Appendix C includes a copy of each of the following STAR data collection instruments:

a. School and System Profile- In order to get an overall picture of each school, principals completed this form which asked for such variables as school enrollment, average daily attendance, average daily membership, Chapter I eligibility, the percentage of students on free lunch, the percentage of students bussed, a breakdown of students by race and total expenditure per student.

b. Principal Profile- provided demographics on the individual principals, i.e., sex, race, education, experience, etc.

c. Teacher Profile- provided background information including college attended and level of education, certification, amount of teaching experience, type of in-service training completed, career ladder level, sex and race.

d. Teacher Log- recorded the time spent on typical daily activities which included routine paper work, student activities, small group, whole group, and individualized instruction, planning and preparation time, and personal time.

e. Grouping Questionnaire- recorded the number of small groups teachers created within their classes for instruction in reading, math, science, and social science, the average number of minutes spent each week in small group instruction and the criteria used for assigning students to instructional groups.

f. Parent/Volunteer/Teacher Interaction Questionnaire- provided the number of times during a four-week period that teachers communicated with parents about the performance or behavior of students or about general classroom activities. Modes of interaction included in person, by phone, or written contact. The quantity and quality of assistance were also noted, including the type of assistance and number of times that assistance was received from a volunteer or BSF aide.

g. Teacher Problem Checklist- indicated the frequency and extent to which teachers were bothered by 61 problems they were likely to encounter. The problems related to their responsibilities to students, their relationships with staff, administrators, and parents, the use of their time, and their professional growth (Cruickshank, D.R., 1980).

h. Exit Interview- Teachers were interviewed "in-person" at the end of the school year. These interviews allowed the teacher to describe differences between teaching in a small class or teaching with a full-time aide and teaching in a regular class. This open-ended interview gave the teachers an opportunity to express their feelings and experiences. The kindergarten interview (1985-86) was unstructured since its primary purposes were to get overall reflections of teachers about their teaching experiences, to thank participants and to serve a public relations function. The interviews provided some important and useful subjective and context data. Based upon an analysis of the first year's interview experience, the researchers developed a more highly structured interview format for subsequent years.

i. Aide Profile- provided information on full-time aides which included education, experience, teaching experience, teaching certification, sex and race.

j. Aide Questionnaire- provided information about an aide's interaction with his/her assigned project STAR teacher. In addition the specific types of daily tasks (e.g., bus duty, lunch duty, teaching lessons, giving tests, etc.) and the amount of time spent on these tasks were reported.

k. Aide Log- provided information about the time full-time aides spent on various generalized categories of activities during a typical day. The activities are the same as the ones described previously under the **Teacher Log** heading.

l. Roster- provided student demographic information such as sex, race, and birthday. Also, at the end of the school year, attendance, promotion, and free-lunch status were reported on the roster.

m. Special Programs Form- identified students who left their classes to participate in special programs such as Chapter I, Special Education, Language Development, Gifted programs, etc. The average amount of time students spent each week in these programs was also recorded.

Based on the data collected in kindergarten, some forms were modified to make them easier to process, and some forms were redesigned in other ways. Some new instruments were devised to collect additional data. These changes did not affect the basic study design but did improve data collection and processing.

Further information regarding testing or data collection instruments may be obtained from the Assistant Commissioner of Curriculum and Instruction, Tennessee State Department of Education, Cordell Hull Building, Fourth Floor North, Nashville, Tennessee 37243-0379.

F. General Description of Key Variables for Analysis

1. Outcome Variables:

a. Stanford Achievement Tests (SESAT II, Primary I, Primary II, Primary III)
b. SCAMIN Self-Concept and Motivation Subscores
c. Promotion/Retention
d. Attendance
e. Teacher Problem Checklist
f. Basic Skills Criterion Tests in Reading & Math (Grades 1,2,3)

Key Variables Cont.

2. **Student Variables:**

   a. Age
   b. Sex
   c. Race
   d. Free Lunch (SES Variable)
   e. Project Entry Date
   f. Total Years in STAR

3. **School Variable**: School Type (Inner City/Rural/Urban/Suburban)

4. **Classroom Variables**:

   a. Class Type (Small/Regular/Regular with Aide)
   b. Average Weekly Volunteer Time
   c. Grouping Practices
   d. Parent-Teacher Interaction

5. **Teacher Variables**:

   a. Teaching Experience (total, at grade level, in this school)
   b. Education Level
   c. Certificates Held
   d. Age (available for K only)
   e. Race
   f. Sex
   g. Instructional Time

6. **Teacher Aide Variables**:

   a. Years of Experience as an Aide
   b. Race
   c. Sex
   d. Age (available for K only)
   e. Use of Time

7. **Comparisons:**

   a. Between/Among Class Conditions and School Types
   b. With Comparison Schools
   c. With Selected State Averages
   d. Within Conditions

G. Methodology (Primary Analysis)

Project STAR's primary analysis consisted of a cross-sectional analysis of data from all students participating in project classes at each grade level, and two longitudinal analyses. For the latter, data were analyzed for students who were in the project in the same class type for four consecutive years (K-I-2-3). Analyses-of-variance procedures were employed to address the major questions of the study as follows:

(1) Class Type (Small/Regular/Aide) was assumed to be a fixed dimension; mean differences among class types comprise the most important question of the investigation.

(2) School Type (Inner City/ Urban/ Suburban/ Rural) was assumed to be a fixed dimension, crossed with class type.

(3) Schools were treated as a random dimension, nested within locations, but crossed with class type, since all three class types were present in each school. This is an important aspect of the design to account for the influence of shared conditions on all project classes within a school.

(4) Classes were treated as a random dimension when there were more than one class of a given type within a particular school.

(5) Students were treated as a random sample, nested within each class. A diagram of the complete design is shown in Figure II-2.

When all of the main effects and interactions of these factors are assembled into an analysis-of-variance model and expected mean squares evaluated, the resulting tests of significance are those given in Table II-4.

TABLE II-4

Analysis of Variance Source Table

Source of Variation	Error Term
Fixed effects:	
Location	Schools
Class Type	Location X Class
Type	Schools X Class Types
Random Effects:	
Schools	
Schools X Class Types	
Classes	
Students	

Figure II-2
A Diagram Of The Project STAR Research Design:
Individual Classes By Class Type Within Schools By Location

LOCATION	Schools	TYPE OF CLASS: SMALL	REGULAR	REGULAR / AIDE
INNER-CITY	1	Class 1,2,3, etc.	Class 1,2,3, etc.	Class 1,2,3, etc.
	2			
	3			
	4			
	5			
	etc.			
SUBURBAN	1	Class 1,2,3, etc.	Class 1,2,3, etc.	Class 1,2,3, etc.
	2			
	3			
	4			
	5			
	etc.			
RURAL	1	Class 1,2,3, etc.	Class 1,2,3, etc.	Class 1,2,3, etc.
	2			
	3			
	4			
	5			
	etc.			
URBAN	1	Class 1,2,3, etc.	Class 1,2,3, etc.	Class 1,2,3, etc.
	2			
	3			
	4			
	5			
	etc.			

Since the error terms needed to test the significance of the fixed effects in the data are variation attributable to Schools and the School-X-Type interaction, student-level data were not required for this portion of the STAR analysis. Thus, data were aggregated to the level of class means before the analysis was conducted, to reduce the magnitude of statistical computations. Table II-4 also shows that the correct error degrees of freedom for the primary questions of the study are proportional to the number of schools -- in the neighborhood of 75 for some tests and 150 for others -- and not the number of students. The exact degrees of freedom for each computer run was affected slightly by the pattern of missing data on the particular instrument.

A parallel analysis was conducted with sex (grades K and 1 only) and race (all grades) as an additional factor of classification. Since both males and females are present in each class and, potentially, both white and minority students, these factors were treated as fixed effects, crossed with all other dimensions in the design. The error terms for treating sex, Sex-X-Location, Sex-X-Type, and Sex-X-Location-X-Type are Schools-X-Sex, Schools-X-Sex-X-Location, Schools-X-Sex-X-Type and Schools-X-Sex-X-Location-X-Type, respectively. For these tests means of all males and all females in each class, or all white and all minority students in each class were used as the units of analysis. Race and Sex were analyzed in parallel computer runs, so that no analysis of both factors simultaneously was conducted.*

The design has unequal **N's** and many empty cells. A general linear model approach for nonorthogonal designs was employed, using the MULTIVARIANCE computer program (Finn and Bock, 1985).

In each year, data from the measurement instruments were analyzed in subsets: the SAT achievement scales, the BSF performance tests (beginning in grade 1), and the SCAMIN self-concept and motivation scales. Since the measures are intercorrelated, multivariate test statistics (Wilks' likelihood ratios) were employed for each subset.

Once a significant main effect or interaction was found, two findings were examined: (A) A univariate test of significance for each scale separately; as follow-up procedure, these are termed "protected" tests; (B) Two orthogonal comparisons among Class Types, when the Class Type effect was found to be significant. The two particular contrasts used were (1) Small Class Means - (Regular + Regular with aide class means)/2 and (2) Regular with aide Class Means - Regular Class Means. The first (1) was selected because no mean differences were found between Regular and Teacher Aide classes in kindergarten, and because children were exchanged between these two Class Types before entering grade 1. The comparison of Small Classes with the average of the other two is not confounded by this procedural modification.

The two contrasts were examined in multivariate form for the entire subset of measures (Hotelling's T^2) and in univariate form for each scale separately (t-tests), but only after an overall test was found to be significant. Again, this is a "protected" procedure. Finally, effect size measures were computed from these contrasts, to reveal the magnitude of the effect (e.g., what impact did reducing the class size really make?).

*This decision was made because means of all white males, minority males, white females and minority females would be based on very small and unreliable groups of youngsters. Also, the magnitude of a combined analysis would be unwieldy.

Prior to all analyses, the distributions of the criterion measures were examined for skewness and outliers. This resulted in only a few deletions of data that were obviously erroneous, and a rescaling of the BSF reading and mathematics scale. Individual students were scored as pass or fail, based on whether or not they passed 80 percent of the objectives covered on the respective test. At the class level, the percentage of students passing each test was obtained (P). Since these were not normally distributed, a "log-odds index" was obtained for each class, ln(P/100-P).

The distribution of the index was normal and used for tests of significance. Descriptive tables in this report, however, give BSF results just as average percent of objectives mastered.

The longitudinal analysis used the same basic design, but in a "repeated measures" form, and with just that subset of students who were in the same experimental condition for three consecutive years. The dependent variables were differences in mean performance between K and grade 1 and between grade 1 and 2; in the second longitudinal analysis, they were differences between grade 1 and 2 and between 2 and 3. Only the SAT measures were scaled as to permit grade-to-grade comparisons of this sort.

The original three years of data are intercorrelated, because they are obtained on the same individuals over time. As a result, the two difference scores are correlated as well. Thus, multivariate repeated measures analyses were used to control statistical errors, in the manner described by Bock (1975). Individual year-to-year growth was examined, or its interaction with other corresponding factors in the design, only when the corresponding overall test was statistically significant.

While the global analyses used the procedures outlined above, other more specific analyses employed a variety of statistical methodologies. These are described in the following chapters of the report, together with the results that were obtained. The analysis procedures employed were conservative and should have provided significant results only when there were considerable differences.