

An Evaluation of Student Achievement in Edison Schools Opened in 1995 and 1996

Gary Miron and Brooks Applegate

The Evaluation Center
Western Michigan University
Kalamazoo, MI 49008
www.wmich.edu/evalctr

December 2000

This evaluation was funded by the National Education Association (NEA), which sought an independent evaluation of schools operated by Edison Schools, Inc. In order to guarantee the evaluators' autonomy, the NEA granted the authors the right to publish and disseminate the report's findings, following delivery of the report to NEA and Edison Schools, Inc. and after allowing "a reasonable time to reply." Both the NEA and Edison Schools, Inc. were invited to comment and provide feedback on drafts of the report throughout the course of the study, although the authors retained responsibility for its content. These groups were also provided copies of this final report at least 10 working days prior to our release of the report. The findings, as well as the views and opinions expressed in this report, are those of the authors and do not necessarily represent the views of the NEA or Edison Schools, Inc.

The Evaluation Center
Western Michigan University
Kalamazoo, MI 49008

2000

Executive Summary

This evaluation’s principal aim was to determine whether or not the Edison model worked—over approximately a 3-year period—to improve student performance on norm- or criterion-referenced standardized tests in the 10 Edison schools that were opened in 1995 and 1996. This report is intended for a wide audience including policymakers, researchers, district personnel, and parents interested in judging the relative merits—and student achievement outcomes—of the first 10 schools opened/operated by Edison Schools Inc.

The intent of our analyses was to incorporate data from a variety of sources and apply a variety of analytical methods to determine the extent of achievement gain made by students enrolled in the schools involved. As an evaluation, we include our own summative judgments by case and across cases. We additionally provide readers with sufficient detail regarding our data types and analytical methods employed so they can make their own judgments regarding the extent of the achievement gains of students in the examined Edison schools. Providing this detail, however, resulted in a rather lengthy and technical report. The Executive Summary, therefore, serves to provide a concise overview of the methods we employed and the results organized by case and across all the cases.

It would seem that addressing student achievement is a rather tangible and noncontestable area to study. This is not the case, however. This study addresses issues that are much on the minds of the public, professional educators, and policymakers. We expect the results to be scrutinized closely. We have stated clearly in the report that we examined student achievement in terms of gains made relative to comparison groups, as opposed to Edison’s preference to evaluate gains made by schools relative to themselves. We don’t contest that the Edison students in this study made gains. The students obviously made gains from year to year, in part due to formal schooling and in part due to nonformal and informal learning activities. On criterion-referenced tests nearly all schools are making gains each year as they adjust their curriculum to state standards. Important questions are whether students in the Edison schools are making gains larger than expected or larger than the district or other relative comparison groups.

Another often heard question specific to Edison—and other for-profit EMOs—is whether the company can “do more for less.” There are two parts to this question: Does Edison do more? And do they do it for less? By focusing on student achievement in schools operated by Edison, this report addresses only the first question. Economists, districts that negotiate contracts with Edison, and even stockholders in the company are in a better position to determine whether Edison operates schools more efficiently than traditional schools. We do not presume that we can answer the financial question. In each of the 10 cases, however, we have included some financial figures based on state and district data. We are aware, however, that these data are limited and that many of the figures are contested by both districts and by Edison.

In terms of inputs, Edison (1999) reports that it is doing more in terms of a longer school day, a longer school year, implementing a rigorous and research-based curriculum, extensive use of technology, etc. Edison also reports that it is doing more in terms of gains made by its students on standardized achievement tests. While this study does not examine what Edison is doing in its schools, and the extent to which its model is implemented, it does examine how students in its first 10 schools did in terms of student achievement.

Recent years have seen a rapid growth in the number of for-profit educational management organizations (EMOs) starting and operating charter schools and competing for contracts to take over operation of poor performing district schools (Miron, 2000). The results from this study contribute to the still limited body of research on the impact and performance of for-profit EMOs.

Description of Edison Schools Inc.¹

Edison identifies three major goals for its schools: (1) move students forward, (2) operate schools for less money, and (3) provide more services. Edison's school model reportedly includes the following: (i) a longer school day and longer school year; (ii) organization of schools based on academies, houses, and teams; (iii) a rich liberal arts curriculum with research-based instruction closely aligned with assessment; (iv) four to six weeks of training for staff prior to the opening of a new school; (v) daily professional development and planning time for teachers; (vi) career ladders for teachers; (vii) a high level of parent and community involvement; (viii) computers and training provided to teachers, students, and families; and (ix) a national system of educational, operational, and financial supports.

Established in 1992 as the Edison Project, the company did not open its first schools until the 1995-96 school year. Edison soon became the largest private educational management organization operating public schools in the U.S. Approximately half of the schools Edison operates are charter schools, while the remaining are schools run under direct contract, or partnership, with local school districts. As of the 2000-01 school year, Edison reports that it is operating 113 schools in 21 states and the District of Columbia, with a combined enrollment of approximately 57,000 students. Edison considers itself to be the first "national system of public schools." Plans for the future assume more expansive growth, including the announcement of new contracts in 2000-01 that include a large commitment in Dallas, Texas, and Inkster, Michigan. All of Edison's initial contracts from 1995 have been renewed except for the one with Sherman Independent School District in Texas where Edison had operated two schools.

Organization of Edison schools. Edison schools are organized by grouping 2 or 3 grade levels into academies. Within the academies, the students are organized into multigrade houses of 100-180 students. The students in each house are largely taught by the same team of teachers throughout the time they are in that academy.

¹ Documentation for this section is derived from Edison's own Web site and from their annual reports (1999 and 2000).

Curriculum design and instructional methods. Edison Schools Inc. has a curriculum that includes reading, math, history/social studies, science, writing, and world language as the core subjects, with classes in character and ethics, physical fitness and health, music, dance, visual art, drama, and practical arts and skills offered at various levels. Four methodological approaches to instruction are reportedly used in the classrooms: project-based learning, direct instruction, cooperative learning, and differentiated learning.

Assessment and accountability. One of the 10 fundamentals of the Edison design recognizes that assessment must provide the foundation for accountability. Edison has developed four main evaluation/assessment areas: (i) state and district tests; (ii) benchmark assessments; (iii) structured portfolios; and (iv) quarterly learning contracts.

Past Studies and Reports

This report summarizes an effort to determine whether or not the Edison model worked to improve student performance on standardized tests. Some attempts have already been made to evaluate overall student achievement at Edison schools. Edison itself has prepared three annual reports on student achievement at its schools (Edison, 1997, 1999, 2000) and has presented general information regarding the gains made at its schools at conferences, workshops, and in the media. The findings contained in Edison's annual reports and the message it spreads in conference presentations and in the media indicate that the company is quite successful and that students enrolled in its schools are making large and substantial achievement gains.

A few evaluations have focused on one or two schools within a given district (Miami Dade, 1998, 1999, 2000; Minneapolis Public Schools, 2000; Wichita State University, 1996) or within a given subject area (see the case studies conducted by Dr. Robert Mislevy between 1996 and 1999). The results from these studies generally have been mixed. In 1998, Dr. Howard Nelson, from the American Federation of Teachers (AFT) prepared a more comprehensive evaluation across most of the then operating Edison schools (AFT, 1998). This study relied more heavily on available state test data and it provided more information on school context than did the Edison reports. The AFT study, while acknowledging that some of the Edison schools were succeeding in raising student test scores relative with comparison groups, found that overall, Edison was not demonstrating gains across its schools. Following the release of the AFT study, Edison employed a Harvard scholar, Dr. Paul Peterson, to reexamine and compare the analyses and results in Edison's own annual reports with the AFT report. While not addressing many of AFT's more critical questions about Edison assessment and reporting activities, Peterson (1998) developed an analytical framework to compare the study designs and found that the designs employed by Edison in its 1997 annual report were stronger than those behind many of the negative or flat trends reported in the 1998 AFT report. The main rationale for this was that more of the trends in the Edison report were based on tracing gains made by individual students, while the trends presented by the AFT relied on consecutive class cohorts (i.e., different groups of students taking the same grade level test from year to year).

In October 2000, the AFT released a second report that provided an overview of the performance of schools operated by Edison. This report relied largely on data from state-mandated criterion-

referenced tests. This report has less narrative to allow for a greater focus on the numbers/data. While this overview still lacked the individual norm-referenced data that Edison also reports on in its annual reports, it did include relevant comparison groups, and the source of the data was clearly marked so others could check the figures for themselves. The overall finding from this study was that “Edison schools mostly do as well or worse than comparable schools; occasionally they do better” (AFT, 2000, p. 6). This finding can be contrasted with the summary of the findings from Edison’s third annual report that was released in September 2000: “For the 1999-2000 school year, the average gain of Edison students, in the core areas of reading, language arts, spelling, writing, and mathematics was 5 percentiles on nationally normed tests and 7 percentage points on criterion-referenced tests, which also include science and social studies. These gains represent improvements of one point in each case over the gains reported for 1995-99, and are the highest gains reported by Edison to date” (Edison, 2000, p. 2).

There are clearly differences in the nature and vested interests of the organizations conducting past studies on Edison. Thus, it is no surprise that they have presented contrasting and contradictory results. Results questioning Edison’s findings have induced strong and critical rebuttals that have, in some cases, been met with equally strong and critical rejoinders.

Aims and Objectives of the Evaluation

This study examined 10 schools operated by Edison Schools Inc. for at least 4 years. We intended to include all 11 schools that opened during the first 2 years of operation (1995-96 and 1996-97). However, we were unable to secure any independently verifiable student achievement data for Dillingham Intermediate School in Sherman, TX. We believe the schools operated by Edison for 4 or 5 years, rather than those open for 3 years or less, provide a more convincing picture of the impact the Edison model can have on student achievement. While some of these schools may have had more difficult start-ups than others, they have been in operation beyond what many consider the start-up phase.

The overriding aim of this study was to examine the impact of Edison schools on student learning as measured by norm- or criterion-referenced tests. In order to achieve this objective, a number of specific tasks were identified:

1. Review and critically assess existing research and evaluations on the impact of Edison schools.
2. Describe the evaluation measures used by Edison Schools Inc.
3. Describe the nature and quality of the standardized test results available.
4. Compare Edison schools in terms of student achievement over time.
5. Compare Edison schools with state and national norms on standardized tests.
6. Compare Edison schools with local school district and state performance levels and—where possible—other similar comparison schools.

7. Develop cases for each of the 10 Edison schools that include (i) a description of the school based upon available literature and documentation, (ii) findings from analysis of norm-referenced and criterion-referenced test results, and (iii) a summary of the diverse results from the analyses of test results.
8. Based upon available literature and documentation, develop a framework for analyzing the 10 cases.
9. Analyze the case studies according to the framework and summarize the results of this analysis.

Sample of Schools and Sources of Information

There is a possibility of selection bias related to the schools selected for the evaluation; therefore, it might be argued that the schools we studied were either performing more poorly or superior to non-selected Edison schools. We examined this possibility by comparing Edison's own school ratings, published in its 2000 annual report, for the 10 schools in this evaluation relative to the remaining 32 schools not considered in this report. Edison rates each school on a 5-point scale, from Strongly Positive to Strongly Negative. There is no indication that the 10 schools we included in this study are rated by Edison any differently from the 32 schools that opened during or after 1997 and for which it reported trend data in its 2000 annual report (Wilcoxon rank-sum test, $p = .7709$). Thus, we believe that while there may be some selection bias in our sample, there is no strong indication that the schools we evaluated are different than the schools for which Edison currently has trend data. Nevertheless, this study is a population study of 10 schools during a specific time period; we do not represent it as an inferential study aimed at all Edison schools and over a more extended time period.

Student achievement data were obtained from a variety of sources. Table 2:1 in Chapter 2 lists the standardized tests in which each participating school took part, and Appendix A describes these tests. Data sets containing individual student results on the norm-referenced tests (ITBS, MAT-7, and SAT-9) were made available to us by Edison Schools Inc. We received 7 such data sets covering 6 of the 10 schools in our study. At the time of our request for data, 3 of these data sets did not contain the results for all possible years, which limited some of these longitudinal analyses. The data sets contained anonymous indicators so that we could trace individual student results over time. From districts and state education agencies we were able to obtain results on the criterion-referenced tests for all 10 schools, the local district, the state, as well as the comparison schools/groups, where applicable.

Description of Our Approach and Unique Attributes of Our Study

Given this context, we are aware that any evaluation of Edison Schools Inc. is going to be carefully scrutinized. This evaluation speaks to both public and private interests in the work of Edison Schools, Inc. and regardless of whether the results are positive or negative, the findings are almost certain to be contested by one group or another. Anticipating this, we employed a number of strategies that we believe have strengthened our analyses as well as our role as an independent evaluator. To

insure our independence from the agency funding the study, the National Education Association (NEA) and the organization being evaluated (Edison Schools Inc.), we have done the following:

- Refused to pursue a larger study including data collection in Edison schools that would require a confidentiality agreement with Edison, which would then control access and release of the results depending on whether or not Edison approved of the findings
- Insisted that our contract with NEA allows us to publish the findings, regardless of the outcomes

In order to strengthen our analyses and the overall credibility of our work, we took the following steps:

- We provided the NEA, Edison Schools Inc., several of the school districts that have Edison schools within their boundaries, and researchers in the field with copies of the draft report and welcomed comments and corrections, which we considered in finalizing the report.
- We sought the advice of experts in the field regarding the strategies and methods we chose to work with state achievement data.
- We sought a variety of data sources and made comparisons with multiple groups whenever possible.
- We clearly describe the assumptions behind our analyses.
- We established and explain the criteria we used to distinguish whether trends are positive, mixed, or negative. We also developed and applied a common method to calculate a mean score across the trends in order to label each school on a 5-point scale—as Edison does—ranging from Strongly Negative to Strongly Positive.
- We attempted to report all data and present all charts and figures clearly and consistently.
- We clearly described the limitations of the study.
- We relied on the program evaluation standards (Joint Committee on Standards for Educational Evaluation, 1994) to guide our work.

To the extent possible, we tried to replicate the presentation of findings that Edison includes in its annual report. This includes a presentation of trends on various norm-referenced or criterion-referenced tests and then a rating given to the school that reflects a summary of all the trends. When we could identify Edison’s own methods and techniques for calculating trends and rating schools, we applied them. When we had to develop our own methods, we explained them as clearly as possible. The following lists some important differences between Edison’s presentation of school results in its annual reports and our own.

- Edison includes only a brief summary of each school, with a table of general data and a page of charts outlining the achievement trends. By contrast, we developed in-depth cases ranging from 10-20 pages in length.
- Edison’s definition of a trend is different than our own. It appears that Edison’s approach to defining and counting trends results in a larger number of trends than our approach. We rate

overall average annual change in a trend, while Edison counts any difference between a two-year block of time as a trend (we explain this in detail in Chapter 2).

- Edison states that it does not intend to compare change in its schools with others (in its annual report there are a number of exceptions to this), while our analysis of gains is dependent on comparisons (see Section 2.3 “To Compare or Not to Compare” for more details).
- Edison reports reviewed contained no information on the number of students (N) reflected in the trends it presents and the source of data is often unclear. By contrast, we made a conscious effort to report the N for all trends and to identify the source of data so that others can check these sources and replicate our work if they wish.
- Edison’s annual reports appeared to us to be intended for a lay audience and provide little insight on the methodology behind the findings they present. Our report is a technical report that provides extensive detail on the methodology, but with an executive summary that presents the results for a lay audience.
- While Edison attempts to present the findings across all its schools, we focused only on those schools that Edison opened in 1995 and 1996 since they have the most data available and because Edison has had time to implement its program more fully in these schools.

Many of the differences between Edison’s reporting and our own are dealt with and explained more thoroughly in other parts of this report.

The conduct of the study was complicated by a number of factors (for example, late receipt of data files from Edison and then only for a limited number of schools, also changes in the members of our evaluation team). The study is marked by a number of other limitations (see Section 2.6). Nevertheless, we are satisfied with the overall results and think this evaluation makes a number of important contributions to the debate over student achievement in schools operated by Edison since it both adds to the record and provides an in-depth discussion and summary of related research on Edison Schools Inc. Also, because of the many years of test results available for our study and because of the use of multiple sources of student achievement data and multiple comparison groups, we believe this is the most in-depth and extensive assessment of student achievement in Edison schools to date.

Methodology

The study combines a variety of statistical methods to compare growth in student achievement in Edison schools with growth in control schools and districts, statewide achievement data, and with national norms, when available. The use of odds ratio analysis to examine criterion-referenced achievement data, given the extensive availability of statewide assessment data and the limitations in the current traditional analyses, may prove to be a significant contribution (Chapter 2 contains more details about odds ratio analysis).

The focus of the methodology employed in this evaluation was to identify and assess the academic achievement gains of students enrolled in Edison schools; therefore, a comparison was essential.

There are a multitude of possible comparisons that could be made, and it is beyond the scope of this report to include them all. Rather, this evaluation focused on two primary types: (1) within-subject gains evidenced by a group of students' year-to-year scores on norm-referenced achievement tests (NRT), and (2) cohort comparisons evidenced by changes across years in the criterion-referenced test (CRT) scores of successive groups of students at given grade levels. These analyses focus on gains over time (i.e., value added) and are based on the assumption that the more exposure to the Edison effect (i.e., the more time a student is enrolled at an Edison school), the better the students will perform compared to relevant comparison groups.

Analysis of Norm-Referenced Achievement Test Data

A variety of NRT data (see Table 2.1 and Appendix A) was provided to us by Edison for 6 of 10 schools. Unfortunately, these data files did not contain data on a comparison sample and tended to be rather incomplete. Consequently, the NRT analyses only examine within-subject change compared with national norms. Following a list-wise deletion strategy, repeated measures ANOVA was used to test for longitudinal trends over the available years. Often Edison provided us with a variety of scores (grade equivalent, standard or scaled score, percentile rank or national percentile rank, and/or the normal curve equivalent). We provide parallel analyses for all types of scores, although gave preference to the normal curve equivalents when summarizing the trends at the end of each case.

Analysis of Criterion-Referenced Achievement Test Data

CRT data were culled from a variety of sources but primarily from the Web (see Table 2:1). The CRT data we examined were based on the aggregate performance of all students taking the test by grade and subject. Thus, unlike the NRT data, where we could follow individual student data over a period of years, the CRT data represented the performance of consecutive cohorts of students. Since these data are available to the public, we were able to define and construct comparison groups (detailed in each case study) for these analyses. The ability to define a comparison group allowed our analyses to test if the relative proportion of students in an Edison school scoring among the various levels coincided with either district or state scoring proportions via a chi-square analysis. These analyses were examined for different years by grade level and subtest category of the state test.

The second analysis strategy we applied to the CRT data examined student learning outcomes as a prospective cohort study by collapsing the distribution of scores on the various state tests into pass/fail categories, reflecting the percent of students meeting or exceeding state standards vs. not meeting state standards. We constructed the 2x2 tables for these analyses to represent the relative odds for a student to fail a component of the state test. The odds ratio (OR) represents the proportion of students who fail the test in the Edison school relative to the proportion of students who fail the test in the comparison school/group.

In order to accomplish the OR analyses, we had to collapse the various CRT scoring categories into pass/fail categories. Although there are several possible ways to do this, we opted to define passing and failing as specified by each state. It should be noted that this reclassification could mask some

important gains evidenced by Edison schools. We acknowledge this possibility and therefore present both the uncollapsed (chi-square) and collapsed (OR) findings so that the reader can determine if there has been a shift in scoring categories that is masked in the collapsed analyses.

We believe this evaluation makes a valuable methodological contribution in its application of a statistical procedure not commonly utilized in educational research or evaluation. With the CRT analyses the use of the Breslow-Day statistic to test for longitudinal trends in consecutive cohorts makes an important methodological contribution that improves evaluators' ability to determine the impact and merit of educational reform over a long period of time. The conventional evaluation approach to CRT data typically involves charting passing/failing trends over time, visual inspection, and quantifying (charting) change scores. The methodological improvement employed in this evaluation was to test for statistically significant differences in passing/failing rates over time. The conventional chi-square often used with this type of data cannot do this, it is limited to examining only one cohort at a time. Thus, in an absolute sense the chi-square tests can identify if the target school is superior to the reference school/group in any given year/grade on any given test; however, it cannot identify any longitudinal gains over several years.

Standards of Comparison

We believe that judgments about the overall performance of the involved schools need to be made on a case-by-case basis. In order to limit potential bias and to establish a common method of making judgments, we thought it was important to establish criteria to distinguish whether or not there had been change over time and whether any identified change was positive or negative. Edison's criteria served as our starting point following which we modified several criteria and added one new criterion. In evaluating trends, we used the following criteria to distinguish when there was meaningful change and when this change was positive or negative:

- *Effect sizes or differences in effect sizes of .20 or greater.* The effect size calculated for the NRT data is the omega squared (ω^2) for a one way repeated measures ANOVA and only provides the reader with an overall effect for time. It does not adequately convey the direction of change nor if the change occurs all in one year or is reflective of a gradual cumulative gain.
- *Differences in national percentile scores of 5 percentage points or more per year*
- *Differences in percentage proficient scores of 5 percentage points or more per year*
- *Differences in grade equivalents of 2 months or more and annual gains in grade equivalents of 14 months or more per year*
- *Differences in DALT gain scores of 2 points or more per year*
- *Differences that are statistically significant (at the .05 level) when tests of significance are available.* The p-value criteria are only applied to the CRT data and not to the NRT data. Utilization of a p-value criteria in the longitudinal NRT analyses does not adequately convey the direction of a statistically significant change. That is, in a longitudinal analysis there can be a statistically significant change in both directions and therefore the p-value is ambiguous relative to the direction of change, only the presence of change. However, in the chi-square and OR

analysis, the p-value conveys a meaningful difference due to the configuration of the contingency tables, in that for these analyses the referent is to the comparison group.

○ *Differences in normal curve equivalents of 3.5 or more per year*

Part of the technical complexity of this report is a function of the variety and large number of analyses conducted. We have constructed summary tables that help guide and focus the reader in distilling the overall impact of Edison in a given school. We treated each analysis category (NRT, CRT) separately. Within each analysis category we rated a finding as negative (-1), mixed (0), or positive (+1) based upon the guidelines presented above.

We also based our rating on a prioritized hierarchy of data. We consider a trend in NRT data to reflect the findings on a group of students as they progress through the life of the analysis by subject and grade level. A trend in CRT data reflects the consecutive cohort findings for a specific grade and subject test over the life of the analysis.² Although we calculated outcomes relative to various comparison groups (e.g., national, state, district, or other), we only counted one trend in the combined table. For NRT data we prioritized the analyses as follows: we considered the normal curve equivalent trend first if available, followed by the percentile rank, then grade equivalent, and lastly the standard score. For CRT data we counted each grade and subject test separately based on the outcomes of the OR Breslow-Day findings relative to the district data.

In its 2000 annual report, Edison defined the 5-point scale it used to rate the overall trends in its schools (Edison assigns one to five stars for each of the categories, from Strongly Negative to Strongly Positive, respectively). Its cut points are as follows: Strongly Negative when 0-19 percent of the trends are positive; Negative when 20-39 percent of the trends are positive; Mixed when 40-59 percent of the trends are positive; Positive when 60-79 percent of the trends are positive; and Strongly Positive when 80-100 percent of the trends are positive.

Since we wanted to consider all the trends and not focus on the positive trends alone, we calculated a mean across the trends where a negative trend is equal to -1, a mixed trend is equal to 0, and a positive trend is equal to +1. We then applied the following 5-point rating scale to the mean trend:

- -1.00 to -0.60 corresponds with “Strongly Negative”
- -0.59 to -0.20 corresponds with “Negative”
- -0.19 to +0.19 corresponds with “Mixed”
- +0.20 to +0.59 corresponds with “Positive”
- +0.60 to +1.00 corresponds with “Strongly Positive”

² By contrast, Edison counts trends in one-year change segments, so a trend of data for a cohort of students over 4 years would be counted by Edison as 3 different trends, while we would count this as one trend and base our rating on the change over the life of the trend.

Methodological Limitations

Several inherent limitations in this evaluation need to be examined in order to provide a balanced interpretation of the findings we have reported and conclusion we have drawn. The limitations to this study can be grouped into three areas—methodology, data quality, and conceptual limitations. They include the following: (i) lack of a comparison group in the longitudinal analyses; (ii) limited and incomplete individual student achievement data supplied by Edison; (iii) validity of comparison groups used in the chi-square and OR analyses; (iv) validity and completeness of Web-based reporting of the district- and state-mandated testing results; (v) evaluation of schools based on student performance alone; and (vi) limited resources. These limitations, which are explained in Chapter 2, should temper all conclusions derived from this evaluation.

Summary of Cases

Roosevelt-Edison Charter School, Colorado Springs, Colorado

Roosevelt-Edison Charter School is a district charter school that was established in 1996 to serve grades K-5. Enrollments in this school totaled 674 during the 1999-00 school year. The school's population was more diverse ethnically than the district and state populations and had a higher proportion of low-income students than the district and state.

Edison provided us a data set that included individual student results on the Iowa Test of Basic Skills. However, this only covered two years (1997-98 and 1998-99). In terms of criterion-referenced tests, we utilized the results from the Colorado Student Assessment Program (CSAP), which is a state-mandated test, based on state standards. We were able to secure test results for the Roosevelt School, the local district, and the state from the Colorado Department of Education. Because the CSAP is a relatively new state assessment program and because it is still not fully implemented in terms of grades and subject tests, we were not able to collect data for all years that the Edison charter school was in operation. We were able to obtain grade 3 reading results for 2 years, grade 4 reading and writing results for 3 years, and grade 5 math results for one year.

Our analyses produced results on nine different trends, six of these based on norm-referenced test results, which examined a longitudinal change in individual student data. Three of the trends were based on the state-mandated assessment test, and we used odds ratio analysis to distinguish if the odds of meeting state standards increased or decreased over time.

Roosevelt-Edison	Positive (+1)	Mixed (0)	Negative (-1)
Norm-Referenced	2 of 6	4 of 6	0 of 6
Criterion-Referenced	0 of 3	2 of 3	1 of 3
TOTALS	2 of 9	6 of 9	1 of 9

This school had a mean trend rating of 0.11. According to our criteria for the 5-point rating scale, the trends in this school are categorized as Mixed. In its 1999 annual report, Edison rated this school as Strongly Positive; and in its 2000 annual report, it rated the achievement gains since opening as Positive and the achievement gains in 1999-00 as Negative.

Our findings are similar to district findings on the District Achievement Level Tests (DALT). Results on the DALT indicate that the school is consistently below the district levels, but the gains made by Roosevelt-Edison on the DALT largely parallel the gains made by the district (i.e., the charter school is not gaining more than the district over time). The results to date indicate that this Edison school does not differ substantially from other district schools.

Henry E.S. Reeves Elementary School, Miami-Dade County, Florida

Henry E.S. Reeves Elementary School is a district contract school serving grades K-5 that Edison began operating in 1996. The total enrollment has remained relatively stable, similar to both the district and the state, with a range from 1,081 students in 1996 to 1,193 students in 1999. Reeves has less diversity ethnically than the state or the district with a much higher proportion of African-American students and lower proportions of white and Hispanic students.

Edison did not provide any NRT data to us for this case, and thus we based our NRT summaries on a previously published evaluation conducted by Miami-Dade County Public Schools (MDCPS) and a dissertation by Sally Shay (2000). Additionally, we were able to extract three years of data on the FCAT and Florida Writes assessments.

MDCPS's ongoing evaluation of Henry E. S. Reeves Elementary School as well as the Shay (2000) study are very likely the most thorough evaluations of any Edison school to date. Because we could not obtain individual results to analyze for this school, we based our NRT trend ratings on results presented by Shay (2000). Her study presented partial effect size estimates (expressed in percentages) from repeated measures analyses. The percentage estimates reflected effect size contributions from both the main effect for group membership and the group by time interaction. Both the district evaluation and the Shay study found that the performance of the Edison students is comparable to but not better than the control groups.

Student results from the three state-mandated tests (FCAT reading and math and Florida Writes) indicate some gains for the Edison students; but in absolute scores, Edison students are still far behind the averages for the district and state. More importantly, the gains made by Reeves on the CRT are similar to those made by the district and state groups.

Henry E. S. Reeves	Positive	Mixed	Negative
Norm-Referenced	0 of 4	4 of 4	0 of 4
Criterion-Referenced	0 of 3	3 of 3	0 of 3
TOTALS	0 of 7	7 of 7	0 of 7

Based on our findings and the discussion of the results, we rated the trends in this school as Mixed with an overall mean trend rating of 0.0. In fact, all seven trends that we considered when summing up this case were Mixed. In its second annual report, Edison (1999) rated the trends on student performance as Positive. In its third annual report, Edison rated the trends for 1999-00 as Strongly Positive and the overall trends since opening as Strongly Positive. Extensive results from the SAT-7 are available for this school. Edison did not consider the findings from the district evaluation, nor from the Shay (2000) study, when presenting the results for this school in its 2000 annual report. In fact, while Edison reported same cohort SAT results for 1996-97 and 1997-98 in its 1999 annual report, it included only limited consecutive cohort SAT results in its 2000 annual report.

Dodge-Edison Elementary School, Wichita, Kansas

Dodge-Edison Elementary School is a district contract school established in 1995 to serve grades K-5. The school's enrollment has ranged from a low of 334 students in 1995 to a high of 657 students during the 1997/1998 school year. Data on the ethnic composition of the students enrolled in the school indicate a majority of white students (65 percent) with little fluctuation over the years. Approximately 58 percent of the students qualify for free or reduced lunches, which is slightly higher than district levels.

Edison provided us with three years of MAT-7 mathematics and reading subtest NRT data from which we were able to extract two cohorts to follow for the years 1995-98 (Cohort A) and 1996-99 (Cohort B). Additionally, we were able to extract a limited amount of data on the Kansas Reading, Math and Writing Assessments, yielding three CRT trends.

Dodge-Edison	Positive	Mixed	Negative
Norm-Referenced	2 of 4	2 of 4	0 of 4
Criterion-Referenced	1 of 3	2 of 3	0 of 3
TOTALS	3 of 7	4 of 7	0 of 7

The average annual gains on the MAT-7 in terms of normal curve equivalents were 5.8 in Math and 4.8 in reading for Cohort A, which were statistically significant. The average annual gains made by Cohort B (2.9 in math and 2.0 in reading), though respectable, were not statistically significant.

The change score for Dodge-Edison students on the Kansas Reading Assessment was more negative than for district students, while the Dodge-Edison students showed a slightly larger gain than the district on Kansas Math Assessment. Differences between Dodge-Edison and the district were small in both of these cases. On the Kansas Writing Assessment, however, the gain made by the Dodge-Edison students was quite substantial and represented 6 percent of the total range of the scale. District students made a very small gain on the Writing Assessment. Therefore, we rated one of three CRT trends as Positive and two as Mixed.

Given the total ratings for the seven trends, which are highlighted above, we rated this school as Positive with a mean trend rating of 0.43. In its 1999 annual report, Edison rated the trends in this school as Positive. In its 2000 annual report Edison rated the achievement gains since opening as Strongly Positive and the achievement gains in 1999-00 as Strongly Positive.

Dodge-Edison has been lauded by Edison as one of its success stories. While the overall label we use to categorize the trends in this school is clearly Positive, they are not as positive as Edison suggests in its third annual report. In any case, the gains made by students enrolled in Dodge-Edison are both substantial and consistent. The achievement gains in this school are clearly the most positive of the 10 cases we examined.

Jardine-Edison Junior Academy, Wichita, Kansas

Jardine-Edison Junior Academy is a district contract school established in 1996 to serve grades 6-8. The school enrolled 822 students during the 1999-2000 school year. Ethnic composition has fluctuated somewhat at this school with the percentage of white students declining from a high of 49 percent during the 1996-97 school year to 32 percent during the 1999-2000 school year. The proportion of students qualifying for free or reduced lunches has fluctuated between 61 and 74 percent, which is consistently higher than district averages, which are around 50 percent.

Edison did not provide us with individual student data for this school; however, we were able to make comparisons between Jardine-Edison and the district on overall performance on the Metropolitan Achievement Test (MAT-7) from data made available by Wichita Public Schools. Additionally, we were able to collect three years of data on the Kansas Reading, Math and Writing Assessments, yielding three CRT trends.

The design behind the trends in the norm-referenced results is based on tracing consecutive cohorts of students at three grade levels over three years. These trends were traced for performance on MAT-7 reading and math. We gave positive ratings to three trends where the Jardine-Edison students made large average annual gains in terms of their national percentile rank as well as large gains compared with districts gains, which were also large, particularly in math.

On the criterion-referenced test results, Jardine had a positive trend relative to the district in reading, but a negative trend relative to the district in writing. Gains made in math by both Jardine-Edison and the district were small and similar in size.

Jardine-Edison	Positive	Mixed	Negative
Norm-Referenced	3 of 6	3 of 6	0 of 6
Criterion-Referenced	1 of 3	1 of 3	1 of 3
TOTALS	4 of 9	4 of 9	1 of 9

Given the total ratings for the 9 trends that are highlighted above, we rated this school as Positive with a mean trend rating of 0.33. In its 1999 annual report, Edison rated the trends in this school as Positive. In its 2000 annual report, Edison rated the trends for the 1999-00 school year alone as Negative, yet at the same time labeled the trends since opening as Strongly Positive. This begs the question or explanation of how the Positive rating of overall trends in 1999 plus the Negative trends for 1999-00 could equal the Strongly Positive rating of overall trends cited in the 2000 annual report (Edison, 2000, p. 78). Although the label we used to categorize the overall trends in this school is Positive, the trends are clearly not Strongly Positive as Edison suggests in its 2000 school report.

Boston Renaissance Charter School, Boston, Massachusetts

The Boston Renaissance Charter School is an independent Charter School established in 1995 to serve grades K-5 with grades 6-8 added in 1996. During the 1998/99 school year there were 1,063 students enrolled in the school. Data on ethnic compositions indicates that the school is predominately nonwhite (86.2 percent), similar to the composition reported for the Boston Public School District; however, Boston Public Schools had more Hispanic and Asian students and fewer African-American students than Boston Renaissance. Both school districts had considerably larger proportions of minority students when compared with the overall state percentage of 77.1 percent white and 32.9 percent nonwhite. This Edison school has a lower proportion of students receiving special education services and a lower proportion of students qualifying for free or reduced lunches than the averages for Boston Public Schools.

Edison provided us with Stanford Achievement Test, Version 9 (SAT9) test data for four years (1996-1999), and this school is also required to take part in the state CRT assessment program. We were able to secure MCAS test results for two years, 1998 and 1999.

Boston Renaissance	Positive	Mixed	Negative
Norm-Referenced	0 of 12	12 of 12	0 of 12
Criterion-Referenced	0 of 6	3 of 6	3 of 6
TOTALS	0 of 18	15 of 18	3 of 18

We examined an abundance of data for this case. We derived 18 overall trends: 12 from NRT data and 6 from CRT data.

We rated this school as Mixed with a mean rating of -0.167. In its 1999 annual report, Edison rated this school as Mixed. In its 2000 annual report, it rated the 1999-2000 school year as Strongly Positive and the achievement gains since opening as Positive.

Seven Hills Charter School, Worcester, Massachusetts

Seven Hills Charter School is an independent charter school established in 1996 to serve grades K-8. The school opened with K-6, and in 1997 grades 7 and 8 were added. During the 1998/99 school

year, 662 students were enrolled in the school. Data on ethnic composition indicate a relatively equal distribution between white and nonwhite students and similar distribution to that found in the Worcester school district. This school has a slightly lower proportion of students with special educational needs than the local school district and has a slightly smaller proportion of students qualifying for free or reduced lunch than the district.

Edison provided us two data sets: Stanford Achievement Test (SAT-9) data for two years (1998-1999) and three years of student achievement data on the Metropolitan Achievement Test (MAT-7) (1997-1999). Similar to all Massachusetts public schools, this school is also required to take part in the Massachusetts Comprehensive Assessment System (MCAS) program. We were able to secure test result data for 1997-98 and 1998-99 only. This matches the CRT results reported by Edison in its 2000 annual school performance report.

Our analyses produced 18 different trends, 12 based on longitudinal norm-referenced tests data and 6 based on consecutive cohort odds ratio analysis of criterion-referenced test data.

Seven Hills	Positive	Mixed	Negative
Norm-Referenced	3 of 12	9 of 12	0 of 12
Criterion-Referenced	0 of 6	1 of 6	5 of 6
TOTALS	3 of 18	10 of 18	5 of 18

This school had a mean trend rating of -0.11 and, according to our 5-point rating scale, this mean score represents a school with Mixed results. In its 1999 annual report, Edison rated this school as Strongly Positive. In its 2000 annual report, it rated the 1999-2000 school year as Mixed and the achievement gains since opening as Mixed also.

It appears that Seven Hills’ students evidenced mild achievement growth over the three academic years on the norm-referenced tests, but did not make any noteworthy gains in comparison with the national norm. The analysis of individual student results, as measured by the MAT-7, indicated largely no gains or losses relative to national norms over the three years we traced the students. Two exceptions to this are when the larger Cohort A showed a statistically significant decrease relative to the national norm for normal curve equivalent (NCE) over two years in language, and the smaller Cohort B showed a statistically significant gain on the NCE over one year in math.

The odds ratio analysis of the MCAS results indicated that the odds of failing (i.e., scoring in the “needs improvement” or “failure” categories) the state assessment test are higher at Seven Hills Charter School than in the district and in the state as a whole. What is most disconcerting is that the odds of Seven Hills’ students failing increases over the two years considered in all subject areas but one in both grades 4 and 8.

In its 1999 annual report, Edison rated this school as Strongly Positive; and in its 2000 annual report it rated the 1999-2000 school year as Mixed and the achievement gains since opening as Mixed also. Our findings, based upon the analysis of individual student results on the MAT-7, SAT-9, and the odds ratio analysis of consecutive cohorts on the MCAS, indicate that the performance of this school—in terms of student achievement—can best be characterized as Mixed with a mean rating of -0.11. This is due to the fact that nearly all the trends were mixed although 3 NRT trends were positive and 5 of the CRT trends were moving in the wrong direction. This is the only school in this study where the rating given by Edison and The Evaluation Center are the same.

Dr. Martin Luther King Jr. Academy, Mt. Clemens, Michigan

Dr. Martin Luther King Jr. Academy (MLK) is a district contract school established in 1995 to serve grades K-5. The school had 877 students enrolled in 1998/99. Data on ethnic composition (for 1997/98) indicates relatively equal distribution between white (53.3 percent) and African-American students (44.3 percent), but scant inclusion of other ethnic groups. The school has around 6 percent of its students receiving special educational services and close to 30 percent qualifying for free or reduced lunches. The proportion of students in the district qualifying for free or reduced lunches has risen in recent years, while the proportion in this and other Edison schools in this district declined.

Edison provided us with a data set containing the Iowa Test of Basic Skills results for three years (1997, 1998, and 1999). Like all Michigan public schools, this school is required to take part in the state CRT assessment program (MEAP). We secured CRT test results for the years 1995-2000 from the Michigan Department of Education for the Edison school, for the only other elementary school in the district with students at grades 4 and 5 (i.e., George Washington Elementary), and for the state.

M.L. King Academy	Positive	Mixed	Negative
Norm-Referenced	0 of 3	3 of 3	0 of 3
Criterion-Referenced	0 of 4	2 of 4	2 of 4
TOTALS	0 of 7	5 of 7	2 of 7

Our analyses produced seven different trends: three based on longitudinal norm-referenced test data and four based on consecutive cohort odds ratio analysis of criterion-referenced test data. Although this represents a moderate number of trends, it is notable that the NRT data represent three-year longitudinal trends and the CRT data represent trends over four or more years.

Three trends were based on the individual student gains on the ITBS, and 4 trends were based on MEAP results. Five of the seven total trends were mixed, and two were negative. Based on these figures, the school had a mean trend rating of -0.29, which indicates an overall rating of Negative. Both the results from the NRT data and the CRT data indicate that students at this school are below national and district norms. This is essentially how the school performed before Edison took over operation. In its 1999 annual report, Edison rated this school as Strongly Positive. In its 2000

annual report, it also rated the 1999-2000 school year as Strongly Positive and the achievement gains since opening as Strongly Positive.

While some gains can be seen in the results from the Iowa Test of Basic Skills, they are not sustained over two years. There were gains in terms of normal curve equivalents (NCE) in the first year and then minor decreases in the NCE in the second year of the longitudinal analysis. One exception to this was in language, where the cohort of students made gradual gains over two years, with the NCE ranking in 1998-99 being statistically significant from the NCE in the 1996-97 school year.

Results from the state MEAP assessment were not promising, however. The odds of not meeting state standards on the subject tests at MLK are generally large. The odds of failing or not meeting state standards for the most part remained the same over time in comparison with the state total and with George Washington Elementary, the only other school in the district with students enrolled in grades 4-5. Both the ITBS and the MEAP results indicated that students at MLK are below national and district norms. This is essentially how the school performed before Edison took over operation.

Mt. Clemens Secondary Academies, Mt. Clemens, Michigan

Mt. Clemens Secondary Academies is a district contract school established in 1996. Both the Edison middle school and the Edison high school share a common building with their district's only other middle and high school. Edison reported that the total enrollment in the secondary academies totaled 481 for the 1999-00 school year. Data on ethnic composition (for 1997-98) indicates slightly more white (62.9 percent) than nonwhite students. According to data presented in Edison's annual reports, the proportion of students receiving special education has dropped in recent years (11.4 percent in 1997-98 and 7.9 percent in 1999-00). The proportion of students qualifying for free or reduced lunches has dropped even more sharply (27 percent in 1997-98 down to 12.5 percent in 1999-00).

Tests administered at the Mt. Clemens academies include the Michigan Educational Assessment Program (MEAP) as well as the Iowa Test of Basic Skills (ITBS). The results from the ITBS were not reported in Edison's second or third annual reports. We were unable to secure individual student achievement data on the ITBS from Edison for this school, although we did receive copies of the summaries of school results for 1996/97 that were supplied to the school by the test company. It was unclear whether or not the school summaries included both the Edison school and the district middle school that share the building.

Given the wide range of grades, this school should have produced a lot of data from the state criterion-referenced test (MEAP). Unfortunately, data were often unusable due to the fact that this school was a school-in-a-school. Thus, we could only secure disaggregated data from a limited number of years. We were able to include MEAP for grades 7 (1997/98 - 1999/00) and 8 (1999/98 and 1999/00), but for grade 11 we were able to examine only the 1998/99 data because instruction at grade 11 began in 1998.

Mt. Clemens Sec. Acad.	Positive	Mixed	Negative
Norm-Referenced	–	–	–
Criterion-Referenced	2 of 8	6 of 8	0 of 8
TOTALS	2 of 8	5 of 8	0 of 8

Because of the limited data available, the school could be grouped with the other schools that Edison opened later and which they categorized as “baseline” rather than attempting to place a label on the progress of the school. If we were to make a judgment based on the limited data available, the Secondary Academies would be rated as Positive with a mean rating of 0.25. In its 1999 annual report, Edison rated this school as Strongly Positive. In its 2000 annual report, it rated the 1999-00 school year as Strongly Positive and the achievement gains since opening as Strongly Positive also.

Mid-Michigan Public School Academy, Lansing, Michigan

Mid-Michigan Public School Academy is a charter school established in 1996 to serve grades K-5, with grades 6-8 added in 1997. This is the largest charter school in Michigan and enrolls more than 1,000 students. Data on ethnic composition indicate a much larger percentage of nonwhite students (79 percent) than is found in Lansing Public Schools. The proportion of students qualifying for free or reduced lunches is approximately 50 percent, which is similar to the district level.

Edison Schools Inc. provided individual student test results on the Metropolitan Achievement Test for one cohort of students between 1997-1999. Like all Michigan public schools, this school is also required to take part in the state CRT assessment program, and we secured test results for the years 1997-1999.

Our analyses produced 11 trends: 3 based on longitudinal norm-referenced tests data and 8 based on consecutive cohort odds ratio analysis of criterion-referenced test data.

Mid-Michigan PSA	Positive	Mixed	Negative
Norm-Referenced	0 of 3	3 of 3	0 of 3
Criterion-Referenced	0 of 8	0 of 8	8 of 8
TOTALS	0 of 11	3 of 11	8 of 11

This school had a mean trend rating of -0.73 and fell into the Strongly Negative category on our 5-point rating scale. Our assessment of this school is similar to the assessment made by the school itself in its 1999 annual report (Mid-Michigan PSA, 1999). Edison’s central office, however, reported a different picture. In its 1999 annual report, Edison rated this school as Mixed. In its 2000 annual report, it rated the 1999-2000 school year as Positive and the achievement gains since opening as Mixed, clearly inconsistent with our findings.

Washington Elementary School, Sherman, Texas

Washington Elementary School is a district contract school that Edison began to operate in 1995 to serve grades K-4. The school had 445 students enrolled during the 1998/99 school year. Data on student characteristics indicate that this school has more minorities and more students qualifying for free or reduced lunches, but fewer students receiving special educational services than other district schools.

Edison did not provide us with individual norm-referenced achievement test data for students in this school, although the ITBS is administered at this school, at least in the earlier grades. However, the Texas Assessment of Academic Skills (TAAS) in reading, math, and writing for Grades 3 and 4, the state-mandated CRT, is administered; and we secured test results for the 1995-96, 1996-97, 1997-98 and 1998-99 school years. The TAAS reporting differs from many state CRTs in that each school is compared with a control group of 40 demographically similar schools, based on the growth in the students' Texas Learning Index³ scores on the TAAS reading and mathematics tests given from one year to the next. The comparable improvement is based only on students whose TAAS results can be matched from the current to the prior year. Since only grades 3 and 4 are tested at Washington, the comparable improvement on the TAAS is based on the performance of fourth graders compared with how they performed on the TAAS in third grade (Texas Learning Index scores are used for comparing individual students from one year to the next).

Based on student performance on the TAAS, we characterize the overall trends in this school as Negative. Except for one successful year, the school is largely unable to even match the gains made in the comparison groups when we consider matched students. In terms of consecutive class cohorts, the school has only been able to come close to matching the control groups on the 4th grade writing test. The results dropped considerably in Edison's first year of operation compared with the previous year when the school was run by the district.

Washington Elementary	Positive	Mixed	Negative
Norm-Referenced	–	–	–
Criterion-Referenced	1 of 5	1 of 5	3 of 5
TOTALS	1 of 5	1 of 5	3 of 5

The proportion of students exempted from the state criterion-referenced test climbed from 4 percent in 1995-96 to around 25 percent in 1998-99, even while the school's performance on the state test remained very poor.

Given the total ratings for the trends, we rate this school as Negative with a mean rating of -0.40. In its 1999 annual report, Edison rated this school as Strongly Positive. In its 2000 annual report, Edison rated the 1999-2000 school year as Positive and the achievement gains since opening as Positive.

³ The Texas Learning Index (TLI) is a score that describes a student's performance on the TAAS reading or mathematics test. It can be used to tell how far a student is above or below the passing standard.

Summary of the Findings

There has been a paucity of evaluations of Edison schools, and those conducted have been typically limited to one or two schools and/or limited in time (Gomez & Shay, 1998, 1999, 2000; Minneapolis Public Schools, 2000; Wichita State University, 1996; AFT, 1998, 2000; Edison, 1997, 1999, 2000). Of particular interest in this evaluation is the relationship of our results to the results previously published in Edison annual reports on student achievement. Thus, we will present a brief overview of findings published in Edison's annual reports (Edison, 1997, 1999, 2000). Edison's First Annual Report on School Performance (Edison, 1997) was structured around five key performance areas: student achievement, customer satisfaction, the implementation of the school design, financial management, and systems growth.

Edison reported that it used the following rules to establish achievement trends: (i) whenever possible to measure trends using the same individual students; (ii) when the same individual students are not available, to measure the same group of students; and (iii) when the same individual students and the same group of students are not available, to measure the same grade level (Edison, 1999). Edison calculated the net gain for every trend from the first observation to the final observation of a trend. Some of its reported trends were three years long, others two years, and some only one.

Edison's second annual report (1999) provided the following summary statements about the achievement levels of its students/schools:

- The overwhelming majority of achievement trends—numbering nearly 200—were positive.
- On average, students were gaining more than 5 percentiles per year against state and national standards.
- Students in nearly every Edison school are achieving more today than when the school opened; in no Edison school are students achieving less.
- Fourteen of 17 Edison schools that have now established achievement trends have moved student achievement forward.
- In total the trends in 10 of the 17 schools were rated Strongly Positive, 4 were rated Positive, 3 were Mixed, and no schools were rated as Negative or Strongly Negative.

The principal aim of this evaluation was to determine whether or not the Edison model worked to improve student performance on norm- or criterion-referenced standardized tests. To that end, this evaluation provides a comparison of *achievement gain*, focusing on two primary types, (1) within-subject gains evidenced by longitudinal panel norm-referenced achievement test (NRT) data and (2) cohort comparisons evidenced by longitudinal cohort criterion-referenced test (CRT) data.

Norm-Referenced Achievement Test Data

Although a variety of NRT data was provided to us by Edison (ITBS, MAT-7 & SAT-9), it was limited in the number of schools and number of years of data collection and did not contain data on

a comparison sample. Consequently, all NRT analyses examined only within-subject change and did not provide an external comparison except for that imbedded in the score reporting scale, e.g., NCE scale. A list-wise deletion strategy, followed by repeated measures ANOVA to test for longitudinal trends over the available years, resulted in cases with complete data over the life of the analysis. However, this procedure likely raised attrition rates. Specific attrition rates can be calculated for each case by cohort group, subject test, and reporting scale and are elaborated in each case and in Chapter 13. Our ratings (positive, mixed, or negative) of NRT trends were based on a system developed and elaborated in Section 2.5.

Overall Results on Norm-Referenced Trends

Case	Positive	Mixed	Negative
Roosevelt-Edison	2 of 6	4 of 6	0 of 6
Henry E. S. Reeves	0 of 4	4 of 4	0 of 4
Dodge-Edison	2 of 4	2 of 4	0 of 4
Jardine-Edison	3 of 6	3 of 6	0 of 6
Boston Renaissance	0 of 12	12 of 12	0 of 12
Seven Hills Charter School	3 of 12	9 of 12	0 of 12
Dr. Martin Luther King Jr.	0 of 3	3 of 3	0 of 3
Mt. Clemens Sec. Academies	–	–	–
Mid-Michigan PSA	0 of 3	3 of 3	0 of 3
Washington Elementary	–	–	–
TOTALS	10 of 50	40 of 50	0 of 50

The main findings from the norm-referenced analyzes can be summarized as follows:

- We charted 50 norm-referenced trends.
- Overall, the norm-referenced trends were either mixed or positive; none were negative.
- Students in Edison schools are generally showing academic achievement gains consistent with grade level advancement on norm-referenced tests.
- Achievement gains do not consistently exceed grade level expectations on norm-referenced tests.

Criterion-Referenced Achievement Test Data

CRT data were culled primarily from the Web. These data typically were reported in aggregate for all students taking the test, broken down by grade and subject test. No individual performance data were obtainable. Thus, the CRT data represented the performance of consecutive cohorts of

students. The criterion-referenced tests were different for each state since they were usually part of a mandated state assessment program. Our ratings (positive, mixed, or negative) were based on a rating system developed and elaborated in Section 2.5. We analyzed the CRT data a number of different ways; however, we based our rating on the consecutive cohort odds ratio (OR) analyses.

A negative rating would be an OR greater than 1.0 with a confidence interval (CI) whose lower bound was greater than 1.0. A mixed rating would be reflective of even odds in an OR analysis, where the CI eclipses 1.0. A positive rating was a protective OR, signified by an OR less than 1.0 with a CI whose upper bound was less than 1.0. With this general scoring system we tried to determine if a trend was present when there were two years of data present by examining the Breslow-Day statistic. In some cases we could not implement these criteria as noted in the table.

Overall Results on Criterion-Referenced Trends

Case	Positive	Mixed	Negative
Roosevelt-Edison	0 of 3	2 of 3	1 of 3
Henry E. S. Reeves	0 of 3	3 of 3	0 of 3
Dodge-Edison	1 of 3	2 of 3	0 of 3
Jardine-Edison	1 of 3	1 of 3	1 of 3
Boston Renaissance	0 of 6	3 of 6	3 of 6
Seven Hills Charter School	0 of 6	1 of 6	5 of 6
Dr. Martin Luther King Jr.	0 of 4	2 of 4	2 of 4
Mt. Clemens Sec. Academies	2 of 8	6 of 8	0 of 8
Mid-Michigan PSA	0 of 8	0 of 8	8 of 8
Washington Elementary	1 of 5	1 of 5	3 of 5
TOTALS	5 of 49	21 of 49	23 of 49

The main findings from the criterion-referenced analyses can be summarized as follows:

- We charted 49 criterion-referenced trends, which is nearly equal to the number of norm-referenced trends.
- Student performance on criterion-referenced tests often lags behind district performance and almost always behind state performance levels.
- In nearly half the trends, we found that students enrolled in Edison schools were making smaller gains on the criterion-referenced tests than comparison groups (i.e., 23 out of 49 trends).
- In 21 out of 49 trends, Edison students showed gains or changes in test results that were similar to the local districts and other comparison groups, and in only 5 of the 49 trends did we find Edison students making larger gains than comparison groups

Overall Findings

Given the varied designs and available data used in this and past evaluations, one can realize that we are far from having a true scientific experiment that can render a conclusive decision on whether or not students succeed in schools operated by Edison Schools Inc. Our results do not differ greatly from past evaluations by Miami-Dade or Minneapolis or those studies conducted by Howard Nelson at AFT (AFT, 1998, 2000). Statewide evaluations of charter schools that included schools operated by Edison also indicate that these schools are not better than other schools (Colorado Department of Education, 2000; Horn & Miron, 1999, 2000; and Renewal Inspection Report, 1999). While there are differences in the quality, scope, and rigor of these studies, it is important to note that the most rigorous studies (Gomez & Shay, 2000; Shay, 2000; and Minneapolis Public Schools, 2000) although limited in scope, all indicated that Edison students were not doing better than comparison groups of students. We are not alone in finding that students in schools operated by Edison—while they often start at levels below national norms and districts averages—progress at rates comparable to students in other district schools. This conclusion indicates that the expectations of district and charter school boards that contract with Edison as well as the expectations of parents who enroll their children in an Edison school are not being met. These groups believe that an apparent goal behind Edison’s school model is to have achievement gains that exceed the gains at comparable schools. Edison advertises this goal in promotional presentations, and the data in its annual reports suggest that it is fulfilling this goal.

Examining all 99 trends in one table, we see an equal proportion of trends from norm- and criterion-referenced analyses. It is important to reflect on our trend definitions. A positive trend reflects students in the Edison school evidencing statistically significant and meaningful annual gains on norm-referenced tests or outperforming the comparison groups on state criterion-referenced tests. A mixed trend is indicative of Edison students gaining at grade-level expectation on norm-referenced tests or performing at the same levels as comparison group students on criterion-referenced tests. A negative trend is evidenced by lower than grade level gains on norm-referenced tests or a lower passing percentage on criterion-referenced tests than the comparison group.

Combined Overall Trends Across All 10 Edison Schools in Our Study

	Positive	Mixed	Negative
Norm Referenced	10 of 50	40 of 50	0 of 50
Criterion Referenced	5 of 49	21 of 49	23 of 49
TOTALS	15 of 99	61 of 99	23 of 99

Our trends for both the norm-referenced tests and the criterion-referenced tests are based on gains in performance and not absolute performance levels. To summarize the main findings in this table, we include the following points:

- The majority of the trends, both norm- and criterion-referenced, were mixed, indicating that students in Edison schools are achieving at levels similar to students in the comparison groups.

- Edison students tend to show larger gains on norm-referenced tests than on state or district criterion-referenced tests. This is likely due to a combination of different factors, two of which are noted: (i) Edison’s curriculum might not be adequately oriented to state standards in the various states in which it works; and (ii) the norm-referenced tests are administered to fewer students, and selectivity may come into play.

Comparison Between Edison’s Findings and The Evaluation Center’s Findings

Our findings suggest that Edison students do not perform as well as Edison claims in its annual reports on student performance. The next table presents a summary of our numerical ratings (5-point scale from strongly positive to strongly negative) for the 10 cases and the ratings Edison gave to these cases in its third annual report. Although this table is based on a number of broad generalizations, it provides a tentative overview of how our findings compare to the findings reported in Edison’s annual performance reports.

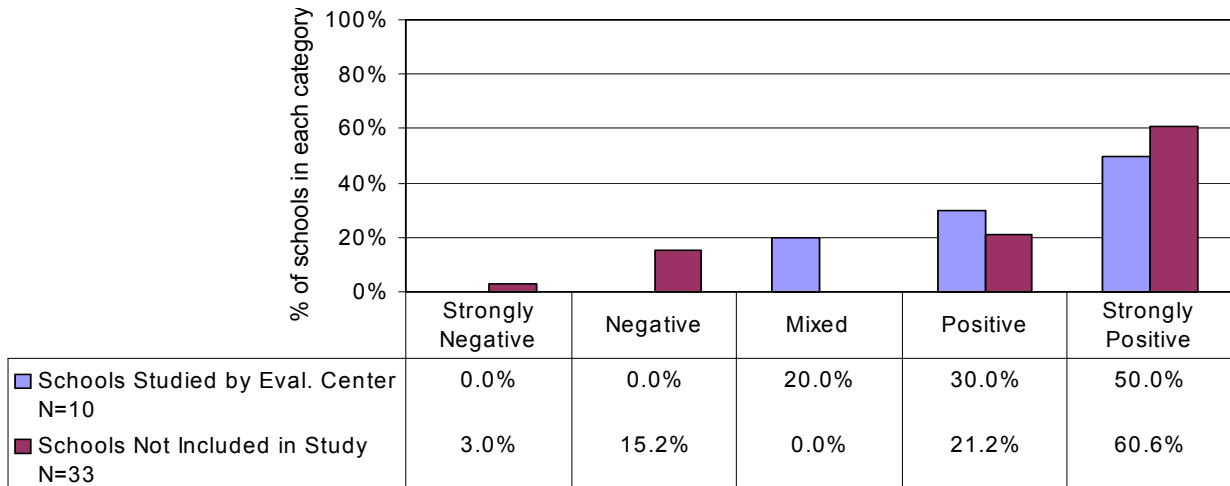
Comparison of Results by Edison Schools and The Evaluation Center

School	Edison’ Schools Inc.				Evaluation Center	
	Edison Rating from 1999 Annual Report	Edison Rating in 2000 for 1999-00 Alone	Edison Rating in 2000 for All Years Since Opening	Numerical Rating	Rating for All Years Since Opening	Numerical Rating
Roosevelt-Edison Charter School	Strongly Positive	Negative	Positive	4	Mixed	3
Henry E.S. Reeves Elementary School	Positive	Strongly Positive	Strongly Positive	5	Mixed	3
Dodge-Edison Elementary	Positive	Strongly Positive	Strongly Positive	5	Positive	4
Jardine-Edison Junior Academy	Positive	Negative	Strongly Positive	5	Positive	4
Boston Renaissance Charter School	Mixed	Strongly Positive	Positive	4	Mixed	3
Seven Hills Charter School	Strongly Positive	Mixed	Mixed	3	Mixed	3
Dr. Martin Luther King Jr. Academy	Strongly Positive	Strongly Positive	Strongly Positive	5	Negative	2
Mt. Clemens Secondary Academies	Strongly Positive	Strongly Positive	Strongly Positive	5	Positive*	4
Mid-Michigan Public School Academy	Mixed	Positive	Mixed	3	Strongly Negative	1
Washington Elementary School	Strongly Positive	Positive	Positive	4	Negative	2

* Indicates very weak basis for rating the school

Average performance rating **4.30** **2.90**

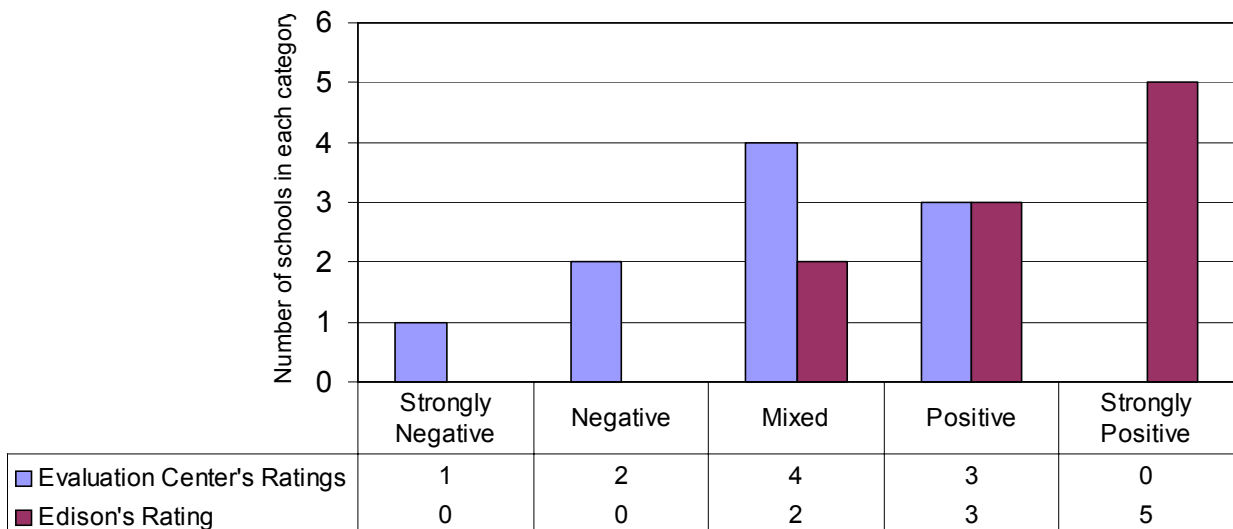
Figure A. Edison’s Own Ratings of the 10 Sampled vs. Nonsampled Schools



In terms of “value-added” performance, over time we found that students at only three of the ten Edison schools were performing better than the comparison groups we examined (overall positive rating). On the other hand, there were three schools whose gains scores were less than those in the comparison groups (overall negative rating), and the remaining four schools showed mixed results.

There is a possibility of selection bias related to the schools included in this evaluation. However, when we analyzed the ratings Edison gave to the 10 schools in this evaluation relative to the 33 nonsampled schools with start-up dates after 1997, we found no significant difference in the ratings Edison published in its third annual report, $p = .7709$; see Figure A. There was, however, a statistically significant difference in the mean ratings we gave the 10 schools based on the findings in this study as compared with Edison’s rating of these same 10 schools, $p < .001$; see Figure B.

Figure B. Ratings of the 10 Sampled Schools by Edison and The Evaluation Center



The differences in Edison’s and our own ratings for the 10 schools included in this study raises an important question. Since disparities exist between Edison’s and our own results for these 10 schools, might they also exist if we were to conduct an analysis of student achievement gains in its other schools?

Recommendations

Several different recommendations are described in detail in Chapter 13. We grouped them into two categories. The first group are general recommendations aimed at improving the general nature of reporting student achievement data in evaluations such as this one. The second group of recommendations is more directly aimed at individuals and groups that are considering contracting out the operation of its schools or already have existing contracts with companies such as Edison Schools Inc.

General Recommendations

There is obviously a large amount of data that a company such as Edison could share with participating schools, districts, and communities as well as with the research community. Listed below are some general recommendations to report student achievement more effectively.

- Prepare more comprehensive and complete annual reports.
- Define trends more clearly and how schools are rated, whether positive or negative.
- Follow the same trends from annual report to annual report.
- Report data across all years with consistent trends.
- Prepare academic reports as well as “lay” reports.
- Report all sample size information with analyses.
- Cite sources for data on schools, particularly when these differ from district and state data.

Specific Recommendations

It is especially important for district and charter school boards to ask questions and make requests in order to benefit from the collaboration with an outside company. Below, we have listed some relative suggestions to consider when contracting out schools. These recommendations are based on the findings presented in the report or reflect lessons learned during the course of the study.

- Require that all student achievement data will be available to the public/researchers.
- Require both external and internal evaluations.
- Obtain and analyze both norm-referenced and criterion-referenced tests
- Ensure that all students are included in the test results and require evaluators to report the number of students included in test results.

- Ensure that all evaluation findings are made available to right-to-know audiences.

Additional recommendations for district or charter school boards when contracting out educational services to EMOs can be found in Lin & Hassel (1999) and Miron (2000).

When seeking new contracts, Edison promises districts and charter school groups that its model is a successful one. In this report, we examined the question of whether Edison did more in terms of student performance on standardized achievement tests. We selected schools that have a long record to trace, and we looked at a variety of test results with its first 10 schools. While our findings do not suggest that Edison did less, they do not suggest that the company did more with these schools in terms of gains on standardized tests.

The differences in Edison's and our own ratings for the 10 schools included in this study raise an important question. Since disparities exist between Edison's and our own results for these 10 schools, might they also exist if we were to analyze student achievement gains in its other schools?

Policymakers and investors want clear and unquestionable results whereas social scientists have a tendency to weigh their findings with cautious interpretations. We attempted to present our findings in a clear and cohesive manner and have tried to alert readers about the many limitations that need to be considered. Nevertheless, we believe that because of the many years of test results examined in our study and because of the multiple sources of student achievement data and multiple comparison groups, this is the most in-depth and extensive evaluation of student achievement in Edison schools to date. We hope that the extensive documentation of our methodology contained in this report will serve those individuals and groups who are interested in looking beyond summary tables and charts and who wish to pursue systematic evaluation of alternative schools.

Table of Contents

Executive Summary	i
Table of Contents	xxix
Lists of Tables, Exhibits, and Appendices	xxxii
List of Acronyms	xxxvii
Acknowledgments	xxxviii
Chapter 1 Introduction and Background to the Study	1
1.1 Introduction	1
1.2 Description of Edison Schools Inc.	4
1.3 Description of Edison’s Evaluation and Assessment Activities	8
1.4 Overview of the Report	9
Chapter 2 Methodology and Related Issues	10
2.1 Aims and Objective of the Evaluation	10
2.2 Sources of Information	11
2.3 To Compare or Not to Compare	13
2.4 Statistical Methods Utilized	16
2.5 Criteria for Evaluating Trends and Summarizing Results for Each Case	22
2.6 Limitations of the Study	25
Chapter 3 Roosevelt-Edison Charter School, Colorado Springs, CO	29
3.1 Descriptive Summary of School	29
3.2 Past Studies and Evaluations	30
3.3 Longitudinal Analysis of Individual Student Data	31
3.4 Chi-Square Analysis of CSAP Data	35
3.5 Odds Ratio Analysis of the CSAP Data	37
3.6 Summary	44
Chapter 4 Henry E. S. Reeves Elementary, Miami-Dade County, FL	46
4.1 Descriptive Summary of School	46
4.2 Past Studies and Evaluations	49
4.3 Summary of Findings by the Miami-Dade County Public School District	49
4.4 Longitudinal Student Outcomes: The Shay Study	52
4.5 Performance on the FCAT and Florida Writes State Assessments	53
4.6 Chi-Square Analysis of FCAT Data	54
4.7 Odds Ratio Analysis of the FCAT Data	58
4.8 Summary	61

Chapter 5	Dodge-Edison Elementary, Wichita, KS	63
5.1	Descriptive Summary of School	63
5.2	Past Studies and Evaluations and Data Available for Analysis	64
5.3	Longitudinal Analysis of Individual Student Data	66
5.4	Comparison with District on the MAT-7	69
5.5	Comparison with District and State on Kansas Assessments	73
5.6	Summary	76
Chapter 6	Jardine-Edison Junior Academy, Wichita, KS	79
6.1	Descriptive Summary of School	79
6.2	Past Studies and Evaluations and Data Available for Analysis	80
6.3	Comparison with District on the MAT-7	80
6.4	Comparison with District and State on Kansas Assessments	84
6.5	Summary	87
Chapter 7	Boston Renaissance Charter School, Boston, MA	90
7.1	Descriptive Summary of School	90
7.2	Past Studies and Evaluations and Data Available for Analysis	91
7.3	Longitudinal Analysis of Individual Student Data	94
7.4	Chi-Square Analysis of MCAS Data	105
7.5	Odds Ratio Analysis of the MCAS Data	107
7.6	Summary	118
Chapter 8	Seven Hills Charter School, Worcester, MA	121
8.1	Descriptive Summary of School	121
8.2	Past Studies and Evaluations and Data Available for Analysis	122
8.3	Longitudinal Analysis of Individual Student Data	123
8.4	Chi-Square Analysis of MCAS Data	129
8.5	Odds Ratio Analysis of the MCAS Data	132
8.6	Summary	143
Chapter 9	Dr. Martin Luther King Jr. Academy, Mount Clemens, MI	147
9.1	Descriptive Summary of School	147
9.2	Past Studies and Evaluation and Data Available for Analysis	148
9.3	Longitudinal Analysis of Individual Student Data	149
9.4	Chi-Square Analysis of MEAP Data	153
9.5	Odds Ratio Analysis of the MEAP Data	156
9.6	Overall Performance on the MEAP Assessments	161
9.7	Summary	167

Chapter 10 Mt. Clemens Secondary Academies, Mt. Clemens, MI	170
10.1 Descriptive Summary of School	170
10.2 Data Available for Our Analyses	172
10.3 Chi-Square Analysis of MEAP Data	173
10.4 Odds Ratio Analysis of the MEAP Data	175
10.5 Overall Performance on the MEAP Assessments	179
10.6 Summary	186
 Chapter 11 Mid-Michigan Public School Academy, Lansing, MI	 188
11.1 Descriptive Summary of School	188
11.2 Past Studies and Evaluations and Currently Available Data	189
11.3 Longitudinal Analysis of Individual Student Data	190
11.4 Chi-Square Analysis of MEAP Data	194
11.5 Odds Ratio Analysis of the MEAP Data	197
11.6 Overall Performance on the MEAP Assessments	204
11.7 Summary	214
 Chapter 12 Washington Elementary School, Sherman, TX	 216
12.1 Descriptive Summary of School	216
12.2 Past Studies and Evaluations and Available Test Data	220
12.3 Analysis of Texas Assessment of Academic Skills Results	221
12.4 Comparable Growth on the TAAS Based on Consecutive Class Cohorts	227
12.5 Comparable Improvement Results Based on Individual Students' Gains on the Texas Learning Index	 231
12.6 Chi-Square Analysis of TAAS Data	234
12.7 Odds Ratio Analysis of the TAAS Data	235
12.8 Summary	243
 Chapter 13 Summary of the Findings	 246
13.1 Edison's Findings	246
13.2 Evaluation Center Findings	247
13.3 Comparison Between Edison's Findings and The Evaluation Center's Findings ..	253
13.4 Recommendations Based on the Findings	255
 References	 257
Appendices	261

List of Tables, Exhibits, and Appendices

List of Tables

Table 2:1	Standardized Test Results from Edison Schools Included in the Study	12
Table 2:2	Construction of 2 x 2 Tables for Odds Ratio Analysis	20
Table 2:3	Types of Analyses Possible for Each School Included in the Study	22
Table 3:1	Table of Sample Sizes for Individual Student Data by Grade and Year	32
Table 3:2	Summary of Chi-Square Findings for Roosevelt-Edison Charter School	36
Table 3:3	Summary of Grade 3 CSAP Odds Ratio Findings for Roosevelt-Edison	38
Table 3:4	Summary of Grade 4 CSAP Odds Ratio Findings for Roosevelt-Edison	38
Table 3:5	Summary of Grade 5 CSAP Odds Ratio Findings for Roosevelt-Edison	43
Table 3:6	Summary of Results for Norm-Referenced Student Achievement Test	44
Table 3:7	Summary of Results on Criterion-Referenced Student Achievement Tests	45
Table 3:8	Combined Overall Trends for Roosevelt-Edison	45
Table 4:1	Distribution of Students by Ethnic Grouping (1997-98)	46
Table 4:2	School, Student and Teacher Background Information for Henry E.S. Reeves Elementary, Miami-Dade, and the State of Florida	48
Table 4:3	Descriptive Statistics for Reeves and Control Groups by Year	53
Table 4:4	Summary of Chi-Square Findings for Henry E. S. Reeves, 1999	58
Table 4:5	Summary of FCAT Odds Ratio Findings for Reeves Elementary School	59
Table 4:6	Summary of Florida Writes Odds Ratio Findings for Reeves Elementary School	59
Table 4:7	Summary Results on Norm-Referenced Tests	61
Table 4:8	Summary Results on Criterion-Referenced Tests	62
Table 4:9	Combined Overall Trends for Henry E.S. Reeves Elementary School	62
Table 5:1	Sample Sizes for Individual Student Data on the MAT-7 by Grade and Year	67
Table 5:2	National Percentile Ranks on MAT-7 Compared with the District	69
Table 5:3	Comparison of Gains in Terms of NPR Between Dodge-Edison and District	70
Table 5:4	Performance on Kansas Assessment Tests Compared with the District and State	73
Table 5:5	Summary of Results on Norm-Referenced Student Achievement Tests	76
Table 5:6	Summary Results on Criterion-Referenced Tests: Comparison of Gains on Kansas Assessments Between Dodge-Edison and Wichita Public Schools	77
Table 5:7	Combined Overall Trends for Dodge-Edison	77
Table 6:1	National Percentile Ranks on MAT-7 Compared with the District	81
Table 6:2	Performance on Kansas Assessment Tests Compared with the District, and State	85
Table 6:3	Comparison of Gains in Terms of NPR Between Jardine-Edison District	87
Table 6:4	Summary Results on Criterion-Referenced Tests: Comparison of Gains on Kansas Assessments Between Jardine-Edison and Wichita Public Schools	88
Table 6:5	Combined Overall Trends for Jardine-Edison	88
Table 7:1	Description of Longitudinal Cohorts	94
Table 7:2	Sample Sizes for Individual Student Data by Grade and Year on the SAT-9	95
Table 7:3	Summary of Chi-Square Findings for Boston Renaissance, Grade 4	106
Table 7:4	Summary of Chi-Square Findings for Boston Renaissance, Grade 8	107
Table 7:5	Summary of Results on Norm-Referenced Student Achievement Tests	118
Table 7:6	Summary Results on Criterion-Referenced Tests	119

Table 7:7	Combined Overall Trends for Boston Renaissance Charter School	120
Table 8:1	Table of Sample Sizes for Individual Student Data on the MAT-7 by Grade and Year	124
Table 8:2	Dependent t-Test Summary Findings for Cohort C and D	129
Table 8:3	Summary of Chi-Square Findings for Seven Hills Charter School, Grade 4	131
Table 8:4	Summary of Chi-Square Findings for Seven Hills Charter School, Grade 8	131
Table 8:5	Summary of Odds Ratio Findings for Seven Hills Charter School, Grade 4	133
Table 8:6	Summary of Odds Ratio Findings for Seven Hills Charter School, Grade 8	134
Table 8:6	Summary of Results on the Metropolitan Achievement Test (MAT-7)	144
Table 8:7	Summary of Results on the Stanford Achievement Test (SAT-9)	144
Table 8:8	Summary Results on MCAS Criterion-Referenced Tests	145
Table 8:9	Combined Overall Trends for Seven Hills Charter School	145
Table 9:1	Table of Sample Sizes for Individual Student Data by Grade and Year on the ITBS	150
Table 9:2	Summary of Chi-Square Findings for MLK, Grade 4	155
Table 9:3	Summaries of the Chi-Square Findings for MLK, Grade 5	156
Table 9:4	Summary of Results on Norm-Referenced Student Achievement Tests	167
Table 9:5	Summary Results on Criterion-Referenced Tests	168
Table 9:6	Combined Overall Trends for Martin Luther King Jr. Academy	169
Table 10:1	Descriptive Information for Mt. Clemens Secondary Academies and District Middle and High Schools	171
Table 10:2	Summary of Chi-Square Findings for Mt. Clemens Academies, Grade 7	174
Table 10:3	Summary of Chi-Square Findings for Mt. Clemens Academies, Grade 8	174
Table 10:4	Summary of Chi-Square Findings for Mt. Clemens Academies, HST	175
Table 10:5	Summary of Odds Ratio Findings for Mt. Clemens Academies, Grade 7	176
Table 10:6	Summary of Odds Ratio Findings for Mt. Clemens Academies, Grade 8	176
Table 10:7	Summary of Odds Ratio Findings for Mt. Clemens Academies, HST	179
Table 10:8	Summary Results on Criterion-Referenced Tests	186
Table 10:9	Overall Summary of Trends	187
Table 11:1	Sample Sizes for Individual Student Data by Grade and Year on the MAT-7	191
Table 11:2	Summary of Chi-Square Findings for Mid-Michigan Academy, Grade 4	195
Table 11:3	Summary of Chi-Square Findings for Mid-Michigan Academy, Grade 5	196
Table 11:4	Summary of Chi-Square Findings for Mid-Michigan Academy, Grade 7	196
Table 11:5	Summary of Chi-Square Findings for Mid-Michigan Academy, Grade 8	197
Table 11:6	Summary of Results on Norm-Referenced Student Achievement Tests	214
Table 11:7	Summary of Criterion-Referenced Test Results for Mid-Michigan	215
Table 11:8	Combined Overall Trends For Mid Michigan PSA	215
Table 12:1	Descriptive Information for Washington Elementary School (1996/97 - 1998/99)	217
Table 12:2	Trends in School Background Characteristics, Washington Elementary School	219
Table 12:3	Washington Elementary Performance on TAAS by Grade and Subject, 1994-1999	222
Table 12:4	Percent Hispanic Students Passing TAAS Compared with Exemptions Because of <i>LEP</i>	228
Table 12:5	Performance on TAAS for All Students Relative to Local District and State 1995-1999	229
Table 12:6	Annual Gains/Loses on the Texas Learning Index Based on Matched Students	232
Table 12:7	Summary of Odds Ratio Findings for Washington Elementary School, Grade 3	236
Table 12:8	Summary of Odds Ratio Findings for Washington Elementary School, Grade 4	236
Table 12:9	Summary of Trends on the State Criterion-Referenced Test, TAAS	244
Table 12:10	Overall Achievement Trends, Washington Elementary School	244
Table 13:1	Approximate Attrition Rates in the Norm-Referenced Test Analysis	249
Table 13:2	Overall Norm-Referenced Trends	250
Table 13:3	Overall Criterion-Referenced Trends	251

Table 13:4 Combined Overall Trends Across All Ten Edison Schools in Our Study	252
Table 13:5 Comparison of Results by Edison Schools and The Evaluation Center	253

List of Exhibits

Exhibit 1:1 Growth of Edison Schools in Terms of Schools and Enrollments	5
Exhibit 3:1 Cohort A at Roosevelt-Edison Charter School: ITBS Results	33
Exhibit 3:2 Cohort B at Roosevelt-Edison Charter School: ITBS Results	34
Exhibit 3:3 Odds Ratio Results for Roosevelt-Edison Charter School	39
Exhibit 3:4 Performance on Grade 3 Reading for Roosevelt-Edison, District, and State	40
Exhibit 3:5 Performance on Grade 4 Reading for Roosevelt-Edison, District, and State	41
Exhibit 3:6 Performance on Grade 4 Writing for Roosevelt-Edison, District, and State	42
Exhibit 3:7 Performance on Grade 5 Math for Roosevelt-Edison, District, and State	43
Exhibit 4:1 Henry Reeves Elementary School, FCAT and Florida Writes Results 1997-1999 Performance on Grade 4 and 5 Tests	55
Exhibit 4:2 Henry E.S. Reeves Elementary School, FCAT and Florida Writes Results 1998-1999 Performance on Grade 4 and 5 Tests	56
Exhibit 4:3 Results of the Odds Ratio Analysis for Henry Reeves Elementary School	60
Exhibit 5:1 Dodge-Edison Results from the Analysis of Individual Students Results	68
Exhibit 5:2 Results on the MAT-7 Reading Between Dodge-Edison and District	71
Exhibit 5:3 Results on the MAT-7 Math Between Dodge-Edison and District	72
Exhibit 5:4 Results on the Kansas Assessments Between Dodge-Edison and District	75
Exhibit 6:1 Results on the MAT-7 Reading Between Jardine-Edison and District	82
Exhibit 6:2 Results on the MAT-7 Math Between Jardine-Edison and District	83
Exhibit 6:3 Results on the Kansas Assessments Between Jardine-Edison and District	86
Exhibit 7:1 Boston Renaissance Cohort A: Individual Student Results on the SAT-9	99
Exhibit 7:2 Boston Renaissance Cohort B: Individual Student Results on the SAT-9	100
Exhibit 7:3 Boston Renaissance Cohort C: Individual Student Results on the SAT-9	101
Exhibit 7:4 Boston Renaissance Cohort D: Individual Student Results on the SAT-9	102
Exhibit 7:5 Boston Renaissance Cohort E: Individual Student Results on the SAT-9	103
Exhibit 7:6 Boston Renaissance Cohort F: Individual Student Results on the SAT-9	104
Exhibit 7:7 Results of the Odds Ratio Analysis for Boston Renaissance (Grade 4)	109
Exhibit 7:8 Results of the Odds Ratio Analysis for Boston Renaissance (Grade 8)	110
Exhibit 7:9 Performance on Grade 4 English Language for Edison, District, and State	112
Exhibit 7:10 Performance on Grade 4 Math for Edison, District, and State	113
Exhibit 7:11 Performance on Grade 4 Science & Technology for Edison, District, and State	114
Exhibit 7:12 Performance on Grade 8 English Language for Edison, District, and State	115
Exhibit 7:13 Performance on Grade 8 Math for Edison, District, and State	116
Exhibit 7:14 Performance on Grade 8 Science & Tech for Edison, District, and State	117
Exhibit 7:15 Performance on Grade 8 History & Social Studies for Edison, District, and State	118
Exhibit 8:1 Seven Hills Charter School: Cohort A: Individual Student Results on the MAT-7	126
Exhibit 8:2 Seven Hills Charter School: Cohort B: Individual Student Results on the MAT-7	127
Exhibit 8:3 Results of Odds Ratio Analysis for Seven Hill Odds, Grade 4	135
Exhibit 8:4 Results of Odds Ratio Analysis for Seven Hill Odds, Grade 8	136
Exhibit 8:5 Performance on Grade 4 English Language for Edison, District, and State	137

Exhibit 8:6	Performance on Grade 4 Math for Edison, District, and State	138
Exhibit 8:7	Performance on Grade 4 Science and Technology for Edison, District, and State	139
Exhibit 8:8	Performance on Grade 8 English Language for Edison, District, and State	140
Exhibit 8:9	Performance on Grade 8 Math for Edison, District, and State	141
Exhibit 8:10	Performance on Grade 8 Science and Technology for Edison, District, and State	142
Exhibit 8:11	Performance on Grade 8 History and Social Studies for Edison, District, and State ...	143
Exhibit 9:1	MLK: Individual Student Results on the ITBS	152
Exhibit 9:2	Grade 4 Odds Ratio Results When Compared to the District and State	158
Exhibit 9:3	Grade 5 Odds Ratio Results When Compared to the District and State	159
Exhibit 9:4	Performance on Grade 4 Math for Edison, District and State	162
Exhibit 9:5	Performance on Grade 4 Reading for Edison, District and State	163
Exhibit 9:6	Performance on Grade 5 Science for Edison, District and State	164
Exhibit 9:7	Performance on Grade 5 Writing for Edison, District and State	165
Exhibit 9:8	Performance on Grade 5 Social Studies for Edison, District and State	166
Exhibit10:1	Results of the Grade 7 Odds Ratio Analysis for Mt. Clemens Secondary Academies ..	177
Exhibit10:2	Results of the Grade 8 Odds Ratio Analysis for Mt. Clemens Secondary Academies ..	178
Exhibit 10:3	Performance on Grade 7 Math for Edison, District and State	180
Exhibit 10:4	Performance on Grade 7 Reading for Edison, District and State	181
Exhibit 10:5	Performance on Grade 8 Science, Writing and Social Studies, 1998-99 and 1999-00 ..	182
Exhibit 10:6	Performance on Grade 8 Science, Writing and Social Studies, 1998-99	183
Exhibit 10:7	Performance on Grade 11 Math and Reading for Edison, District, and State	184
Exhibit 10:8	Performance on Grade 11 Science and Writing for Edison, District, and State	185
Exhibit 11:1	Mid Michigan PSA: Analysis of Individual Student Results on the MAT-7	193
Exhibit 11:2	Results of the Odds Ratio Analysis Grade 4 Reading and Math	198
Exhibit 11:3	Results of the Odds Ratio Analysis Grade 7 Reading and Math	199
Exhibit 11:4	Results of the Odds Ratio Analysis Grade 5 Science and Writing	200
Exhibit 11:5	Results of the Odds Ratio Analysis Grade 8 Science and Writing	201
Exhibit 11:6	Grade 4 Reading MEAP Results for Mid Michigan, District, and State	205
Exhibit 11:7	Grade 7 Reading MEAP Results for Mid Michigan, District, and State	206
Exhibit 11:8	Grade 4 Math MEAP Results for Mid Michigan, District, and State	207
Exhibit 11:9	Grade 7 Math MEAP Results for Mid Michigan, District, and State	208
Exhibit 11:10	Grade 5 Science MEAP Results for Mid Michigan, District, and State	209
Exhibit 11:11	Grade 8 Science MEAP Results for Mid Michigan, District, and State	210
Exhibit 11:12	Grade 5 Writing MEAP Results for Mid Michigan, District, and State	211
Exhibit 11:13	Grade 8 Writing MEAP Results for Mid Michigan, District, and State	212
Exhibit 11:14	Grades 5 & 8 Social Studies: MEAP Results for Mid Michigan, District, and State ...	213
Exhibit 12:1	Washington Elementary School TAAS Results 1994-1999 Grade 3 Comparisons Reading and Math	224
Exhibit 12:2	Washington Elementary School TAAS Results 1994-1999 Grade 4 Comparisons Reading and Math	225
Exhibit 12:3	Washington Elementary School TAAS Results 1994-1999 Grade 4 Comparisons Writing and All Tests	226
Exhibit 12:4	Average Gains/Losses on the Texas Learning Index, Washington Elementary School	233
Exhibit 12:5	Results of the Odds Ratio Analysis for Washington Elementary Grade 3 vs. Campus ..	237
Exhibit 12:6	Results of the Odds Ratio Analysis for Washington Elementary Grade 3 vs. District ..	238
Exhibit 12:7	Results of the Odds Ratio Analysis for Washington Elementary Grade 3 vs. State	239
Exhibit 12:8	Results of the Odds Ratio Analysis for Washington Elementary Grade 4 vs. Campus ..	240
Exhibit 12:9	Results of the Odds Ratio Analysis for Washington Elementary Grade 4 vs. District ..	241

Exhibit 12:10 Results of the Odds Ratio Analysis for Washington Elementary Grade 4 vs. State	242
Exhibit 13:1 Edison’s Own Ratings of the Sampled Schools vs. Nonsampled Schools	254
Exhibit 13:2 Edison’s Own Ratings As Compared to The Evaluation Center Ratings For the Ten Schools Considered in This Study	255

List of Appendices

Appendix A Description of Norm- and Criterion-Referenced Tests Considered in the Study . . .	261
Appendix B Normal Curve Equivalents Based on National Percentile Ranks	267
Appendix C Criteria for Rating Student Achievement Trends Presented by Edison and AFT . .	268
Appendix D Results Provided by Edison for Dillingham Intermediate School, Sherman, TX . .	270
Appendix E Chi-Square Results for All Schools	271

List of Acronyms

AEIS	Academic Excellence Indicator System	MAT-7	Metropolitan Achievement Test, 7 th Edition
AFT	American Federation of Teachers	MCAS	Massachusetts Comprehensive Assessment System
ANOVA	Analysis of variance	MDCPS	Miami-Dade County Public Schools
BRCS	Boston Renaissance Charter School	MDE	Michigan Department of Education
B-D	Breslow-Day statistic	MDOE	Massachusetts Department of Education
CDE	Colorado Department of Education	MEAP	Michigan Educational Assessment Program
CI	Confidence interval	MLK	Dr. Martin Luther King Jr.
CRT	Criterion-Referenced Test	NEA	National Education Association
CSAP	Colorado Student Assessment Program	NCE	Normal curve equivalent
CTBS	Comprehensive Test of Basic Skills	NPR	National Percentile Rank
DALT	District Achievement Level Tests	NRT	Norm-Referenced Test
DEPS	Dodge-Edison Partnership School	OR	Odds ratio
EMO	Educational management organization	PR	Percentile rank
EC	The Evaluation Center, WMU	PSA	Public School Academy
ES	Effect Size	SAT-9	Stanford Achievement Test, Version 9
ESL	English as a second language	SES	Socioeconomic status
FCAT	Florida Comprehensive Assessment Test	SFA	Success for All
FDE	Florida Department of Education	SPSS	Statistical Package for the Social Sciences
GE	Grade equivalent	SS	Standard score
HLM	Hierarchical linear model	Stn	stanine scores
HST	High School Test	TAAS	Texas Assessment of Academic Skills
ISD	Intermediate School District	TEA	Texas Education Agency
ITBS	Iowa Test of Basic Skills	TLI	Texas Learning Index
KDE	Kansas State Department of Education	WMU	Western Michigan University
LEP	Limited English proficiency		

Acknowledgments

This was a complex study, and it could not have been completed without the cooperation and commitment of a number of individuals. First and foremost, we want to thank the National Education Association for providing some of the funds for this important study. At the NEA, Heidi Steffens and Phil McLaurin, from the Office for Public Education Advocacy, provided oversight and guidance in the conduct of the study. Anthony Rolle, formerly a statistician at NEA, also contributed ideas and comments on the statistical methods employed. We also wish to thank Edison Schools Inc. for sharing with us individual data sets for 6 of the 10 schools in our study. John Chubb, Chief Education Officer at Edison, exchanged thoughts and ideas with us during various phases of the work; and he and Christopher Cerf, the Chief Operating Officer at Edison, provided us with general comments on draft chapters of the report.

We would also like to thank the various school districts that provided documentation and data and in many cases also reviewed and commented on draft cases. Likewise, the study benefited from a number of statisticians and researchers who provided suggestions on methods and design as well as feedback on some of the chapters.

We were fortunate to have a talented evaluation team at our side and a number of research assistants who helped us with a wide variety of tasks. Carla Howe, Deborah Lehmann, and Kim Reynolds were instrumental in helping us organize and summarize data from numerous annual school reports, Web sites, and a variety of district and state documents. They also assisted by reviewing drafts of many of the chapters. In addition, student assistants Santos Kalaskar and Archana Prakash provided data entry support. We were also fortunate to have expert assistance from Christopher Nelson, who reviewed and provided comments on some of the chapters, and Sally Veeder, who helped with editing. Finally, we express our appreciation to Daniel Stufflebeam, the director of The Evaluation Center, and Arlen Gullickson, Chief of Staff, who provided guidance in the overall organization and design of the evaluation.

From this long list of names and contributions, one can clearly see that this was a team effort. We are confident that the evaluation has benefited from the expertise and diverse perspectives of the various contributors and evaluation team members. While we recognize and express our appreciation for the contributions made by these many persons, we are mindful that we are responsible for the content of the report, including errata.

Gary Miron and Brooks Applegate
December 20, 2000