Research and Evaluation into the Curry/Samara Model Of Curriculum, Instruction, and Assessment.

by

Michael L. Connell, Ph.D.
University of Houston
Curriculum and Instruction
CUIN 5872
Houston, TX 77069-5872
Mkahnl@aol.com

REPORT
11/04/2003

Prepared in cooperation with
The Aldine Independent School District and
The Curriculum Project

Table of Contents

Appendix B & C are omitted from this PDF file to decrease memory requirements, but are available upon request, please call 800.867.9067 for more information.

Introduction and Overview

The research presented in this report addresses two foundational areas concerning the *Curry/Samara Model of Curriculum, Instruction, and Assessment*. The first may be classified as providing empirical, replicable and quantitative evidence that adoption and use of the Curry/Samara Model enables positive impacts upon student classroom performance. The second, and in many ways more interesting of these areas, is an attempt to describe some of the positive benefits accruing from continuing use of the Curry/Samara Model over time - particularly as reflected in long-term patterns of student achievement and curriculum implementation.

For this report two parallel elementary schools were carefully selected from within Aldine Independent School District (AISD)- one of which (Control Elementary[1]) was described by district administrators and personnel as not significantly participating in the Curry/Samara Model and the other (Treatment Elementary) which was classified as participating at a high level. These schools served as the basis for both strands of the analysis.

The primary research concerns were addressed via a two-part research and evaluation scheme involving two mutually supporting research methodologies - Analysis of Variance (ANOVA) and Student-Problem Analysis (S-P Chart). The use of the

[1] This report has modified the actual school names to ensure confidentiality as requested by Aldine Independent School District.

Analysis of Variance methodology allows statistically significant claims to be made regarding observable differences between these two groups as reflected in their student performance scores. The use of the Student-Problem Chart allows in-depth investigation of differences in student understanding, as measured by consistency of item response, and objective mastery - as well as providing insight into the degree to which teachers were able to successfully align their instruction to the objectives evaluated. Together these two methodologies allow for comprehensive A vs. B comparisons to be made and provide a series of pedagogical and curricular insights into the nature of the differences observed - addressing the identified research concerns.

Populations and Data Sources

The schools chosen for participation in this study were carefully chosen to match on a number of important variables including gender, Title 1A status, educational "at-risk" classification, etc. Although in educational research a perfect match is rarely possible, these schools were found to be comparable in most areas. Furthermore, in several areas such as number of students classified as being educationally at-risk and number of students classified as limited English proficient (LEP), the control school was initially in a more favorable position. A series of frequency charts providing matching demographics of these schools is provided in Appendix A.

Data on student mastery from the state mandated end of year *Texas Assessment of Academic Skills* (TAAS) for 1997, 1998, 1999, 2000, 2001, and 2002 were used for each

analysis[2]. The data provided by the district was in its raw form as directly received from the state ensuring its integrity. From this raw data a Statistical Package for the Social Sciences (SPSS) file was constructed for each year, mapping each element of the data line vector with their appropriate variables and labels, as described in the document Data File Format with Student Item Analysis[3]. These six data sets were then used as the source for demographic analysis and school selection, as well as the Analysis of Variance studies, and Student Problem Chart development.

A-B Comparisons

Findings and Discussion

To address the first research area, concerning empirical, replicable and quantitative evidence supporting adoption and use of the Curry/Samara Model, a series of Analysis of Variance (ANOVA) studies were conducted. These ANOVA studies compare student data by grade, on the number of TAAS objectives mastered each year in the core content areas of reading, writing, and mathematics. This data was drawn from two parallel elementary schools - one of which was identified as not significantly participating in the Curry/Samara Model (control) while the other was classified by district administrators as participating at a high level (treatment).

This section will now report each of the parallel series of Analysis of Variance for the years 1997, 1998, 1999, 2000, 2001, and 2002. In these analyses, and all subsequent plots, the number of objectives mastered in the area of **Writing** is indicated by

---

[2] A special thanks is offered to Area Superintendent Margarita Byrum and Monica Uphoff, administrative assistant of data analysis for their help and support in providing the six years of TAAS data utilized in this report.

[3] http://www.tea.state.tx.us/student.assessment/reporting/formats/

W_#MAST, the number of objectives mastered in the area of **Reading** is indicated by **R_#MAST**, and the number of objectives mastered in the area of **Mathematics** is indicated by **M_#MAST.**

1997 ANOVA of Core TAAS Objectives

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval fo Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	110	2.09	1.246	.119	1.86	2.33	0	3
	Treatment	114	1.85	1.459	.137	1.58	2.12	0	3
	Total	224	1.97	1.361	.091	1.79	2.15	0	3
R_#_MAST	Control	225	3.71	2.383	.159	3.40	4.02	0	6
	Treatment	243	3.70	2.772	.178	3.35	4.05	0	6
	Total	468	3.70	2.590	.120	3.47	3.94	0	6
M_#_MAST	Control	225	7.924	4.0925	.2728	7.387	8.462	.0	12
	Treatment	243	7.597	5.2075	.3341	6.939	8.255	.0	12
	Total	468	7.754	4.7024	.2174	7.327	8.181	.0	12

ANOVA

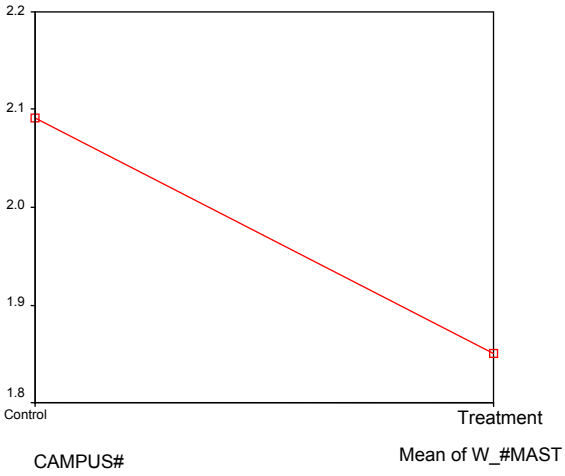
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	3.225	1	3.225	1.748	.187
	Within Groups	409.556	222	1.845		
	Total	412.781	223			
R_#_MAST	Between Groups	.029	1	.029	.004	.948
	Within Groups	3131.687	466	6.720		
	Total	3131.716	467			
M_#_MAST	Between Groups	12.549	1	12.549	.567	.452
	Within Groups	10314.193	466	22.133		
	Total	10326.741	467			

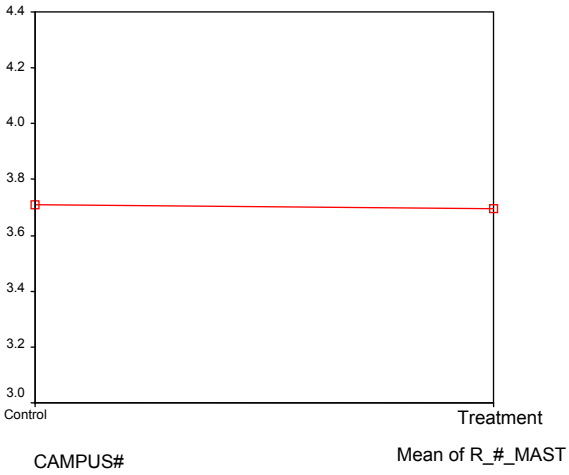
It should be noted that although none of these observed differences in student mastery of objectives are statistically significant, in each case the control school is scoring slightly better. This advantage at the time of the initial data would certainly lend credence to the argument that any future growth in student performance could be attributed to the continuing implementation and use of the Curry/Samara Model. At a

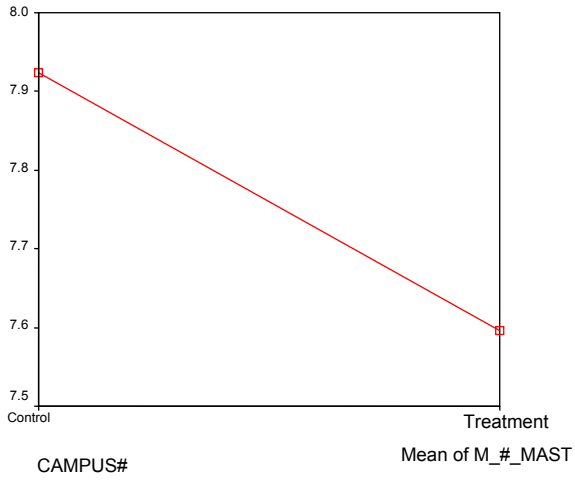
bare minimum it illustrates the beginning comparability of the two schools chosen for use in this study.

The following means plots illustrate these observed differences.

1997 Means Plots







These graphs clearly show that at the time the baseline for this research was drawn there was not an unfair advantage given to the treatment group by selection of the schools. Indeed, the treatment school is at a noticeable disadvantage in each area measured. As earlier indicated this is also reflected in the number of students classified as at-risk and Limited English Proficient (LEP).

When one considers the initial disadvantages demonstrated by the student performance at the treatment school it makes the later growth illustrated in the following ANOVA's all the more impressive.

1998 ANOVA of Core TAAS Objectives

Descriptives

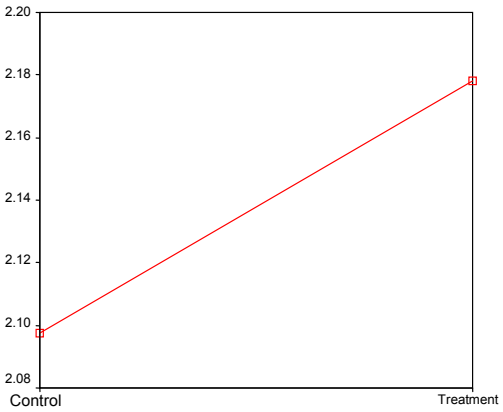
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	113	2.10	1.172	.110	1.88	2.32	0	3
	Treatment	146	2.18	1.258	.104	1.97	2.38	0	3
	Total	259	2.14	1.220	.076	1.99	2.29	0	3
R_#_MAST	Control	228	4.20	1.974	.131	3.94	4.46	0	6
	Treatment	271	4.60	2.309	.140	4.33	4.88	0	6
	Total	499	4.42	2.169	.097	4.23	4.61	0	6
M_#_MAST	Control	228	8.689	3.4302	.2272	8.241	9.136	.0	12
	Treatment	271	9.192	4.2470	.2580	8.684	9.700	.0	12
	Total	499	8.962	3.8994	.1746	8.619	9.305	.0	12

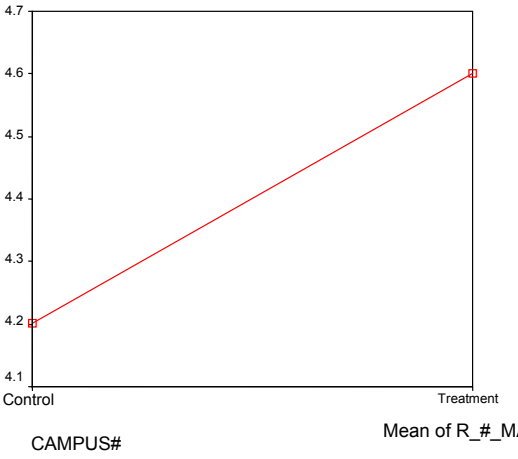
ANOVA

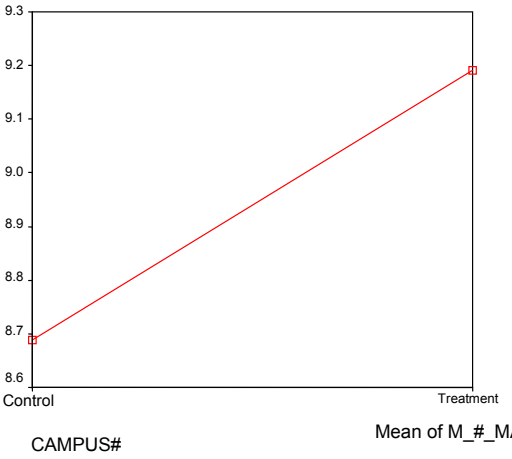
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	.415	1	.415	.278	.598
	Within Groups	383.299	257	1.491		
	Total	383.714	258			
R_#_MAST	Between Groups	19.784	1	19.784	4.232	.040
	Within Groups	2323.679	497	4.675		
	Total	2343.463	498			
M_#_MAST	Between Groups	31.364	1	31.364	2.067	.151
	Within Groups	7540.912	497	15.173		
	Total	7572.277	498			

By 1998 the baseline situation presented in the previous ANOVA has been significantly changed to the benefit of the treatment school. Although these results are only statistically significant in the case of reading ($p \leq .04$), it shows a marked turnaround from the baseline. This difference is made very clear when examining the means plots which follow.

1998 Means Plots







A visual inspection of the slope of these plots when compared with the previous year provides for a striking illustration of the improvement measured. Although statistical significance has yet to be achieved in each content area, the impact of the Curry/Samara Model can already be seen on student improvement as measured by the TAAS.

As we will show in the subsequent analyses not only will significant statistical significance be achieved, but at a level rarely seen in educational research.

1999 ANOVA of Core TAAS Objectives

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	116	2.00	1.292	.120	1.76	2.24	0	3
	Treatment	141	2.11	1.302	.110	1.89	2.32	0	3
	Total	257	2.06	1.296	.081	1.90	2.22	0	3
R_#_MAST	Control	245	3.96	2.453	.157	3.65	4.26	0	6
	Treatment	269	4.50	2.304	.141	4.23	4.78	0	6
	Total	514	4.24	2.390	.105	4.03	4.45	0	6
M_#_MAST	Control	245	7.690	4.3014	.2748	7.149	8.231	.0	12
	Treatment	269	9.405	3.7620	.2294	8.954	9.857	.0	12
	Total	514	8.588	4.1145	.1815	8.231	8.944	.0	12

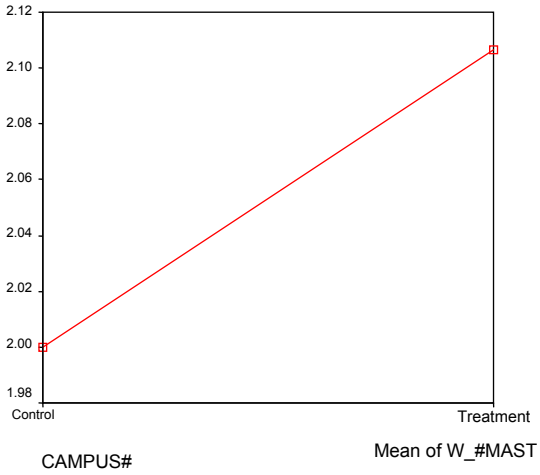
ANOVA

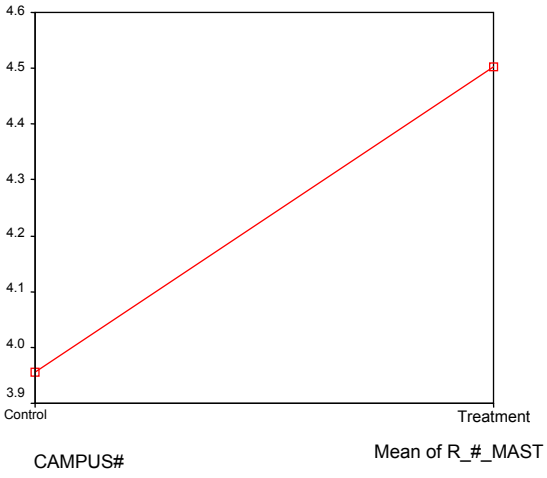
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	.720	1	.720	.428	.514
	Within Groups	429.404	255	1.684		
	Total	430.125	256			
R_#_MAST	Between Groups	38.330	1	38.330	6.787	.009
	Within Groups	2891.755	512	5.648		
	Total	2930.086	513			
M_#_MAST	Between Groups	377.303	1	377.303	23.254	.000
	Within Groups	8307.257	512	16.225		
	Total	8684.560	513			

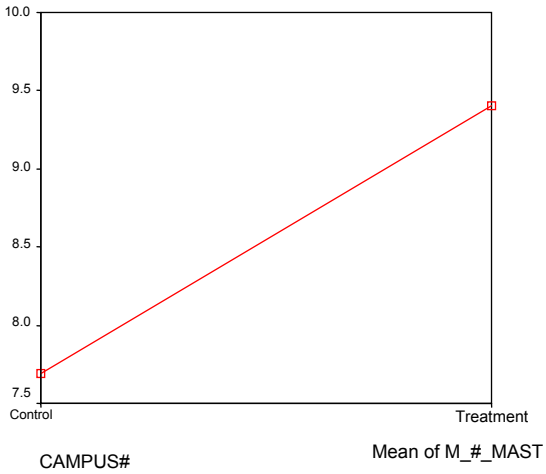
By the third year student performance at the treatment school has reached statistical significance in the core content areas of reading and mathematics. Furthermore, these results go far beyond the levels typically associated with educational significance ($p \leq .05$). For the case of reading we see that the probability of this difference being due to chance is ($p \leq .009$) and for mathematics ($p \leq .0005$). To find significance at this level is more typical of medical research than educational or social research.

The following mean plots illustrate this continuing, and now in two cases statistically significant, growth in student achievement on the part of the treatment school.

1999 Means Plots






10.0
9.5
9.0
8.5
8.0
7.5
Control
Treatment
CAMPUS#
Mean of M_#_MAST

2000 ANOVA of Core TAAS Objectives

Descriptives

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval fo Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	137	1.90	1.190	.102	1.70	2.10	0	3
	Treatment	134	2.41	1.028	.089	2.23	2.59	0	3
	Total	271	2.15	1.140	.069	2.01	2.29	0	3
R_#_MAST	Control	277	4.16	2.347	.141	3.88	4.44	0	6
	Treatment	269	5.00	1.833	.112	4.78	5.22	0	6
	Total	546	4.58	2.148	.092	4.39	4.76	0	6
M_#_MAST	Control	277	7.957	4.0973	.2462	7.472	8.441	.0	12
	Treatment	269	9.870	3.1412	.1915	9.493	10.247	.0	12
	Total	546	8.899	3.7776	.1617	8.582	9.217	.0	12

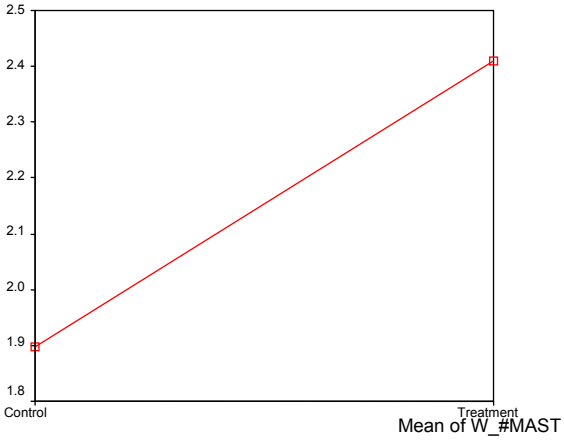
ANOVA

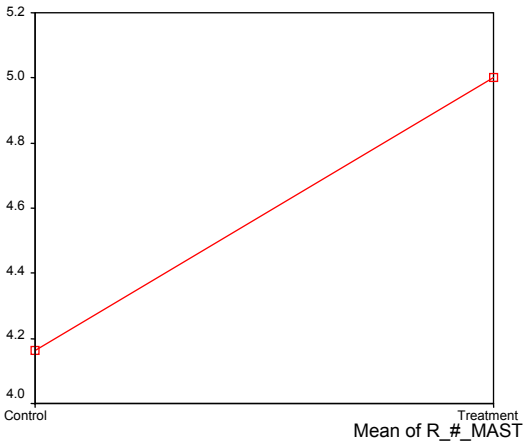
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	17.802	1	17.802	14.381	.000
	Within Groups	332.995	269	1.238		
	Total	350.797	270			
R_#_MAST	Between Groups	95.732	1	95.732	21.523	.000
	Within Groups	2419.690	544	4.448		
	Total	2515.421	545			
M_#_MAST	Between Groups	499.533	1	499.533	37.338	.000
	Within Groups	7277.926	544	13.379		
	Total	7777.460	545			

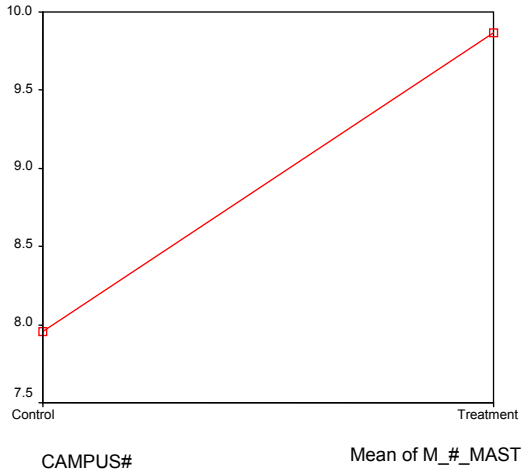
By the fourth year the superior student performance at the treatment school has resulted in the achievement of statistical significance in all three areas of writing, reading and mathematics. Furthermore, these results again go far beyond the levels typically associated with educational significance (in this case $p \leq .0005$). Given the initial inequities between the two schools in favor of the control group, to have achieved such amazingly significant results is truly impressive.

The mean plots illustrate this now three-way statistically significant growth in student achievement.

2000 Means Plots







2001 ANOVA of Core TAAS Objectives

Descriptives

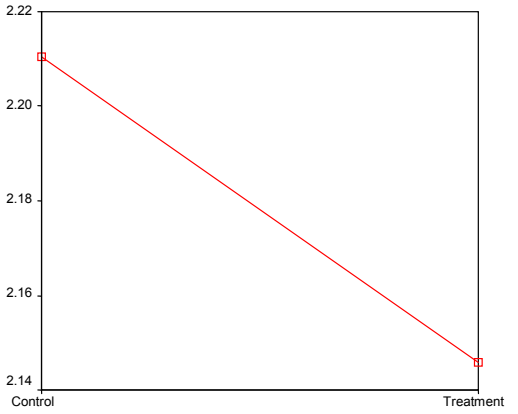
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval fo Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	133	2.21	.922	.080	2.05	2.37	0	3
	Treatment	151	2.15	.969	.079	1.99	2.30	0	3
	Total	284	2.18	.946	.056	2.07	2.29	0	3
R_#_MAST	Control	280	4.40	1.775	.106	4.19	4.61	0	6
	Treatment	266	5.04	1.633	.100	4.84	5.24	0	6
	Total	546	4.71	1.736	.074	4.57	4.86	0	6
M_#_MAST	Control	280	7.211	2.9625	.1770	6.862	7.559	.0	12
	Treatment	266	9.729	1.8208	.1116	9.510	9.949	.0	12
	Total	546	8.438	2.7735	.1187	8.205	8.671	.0	12

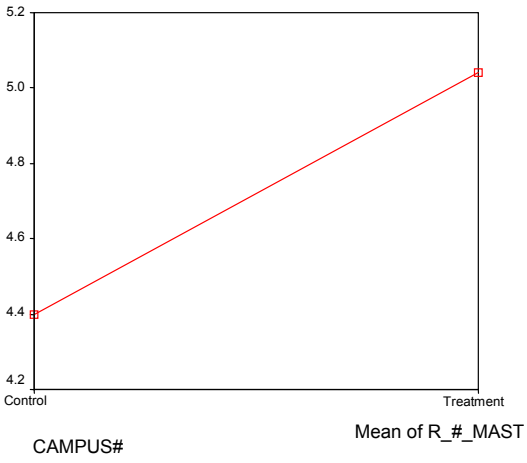
ANOVA

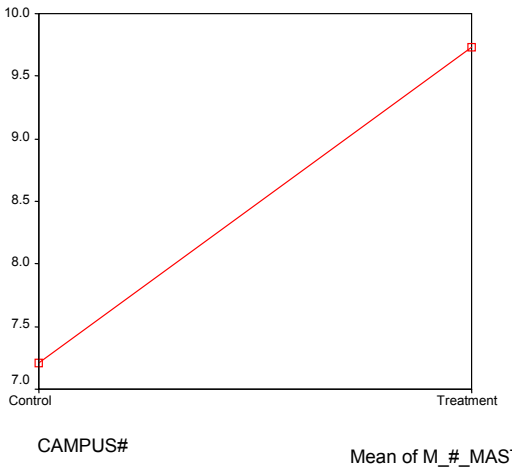
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	.297	1	.297	.331	.565
	Within Groups	252.900	282	.897		
	Total	253.197	283			
R_#_MAST	Between Groups	56.110	1	56.110	19.249	.000
	Within Groups	1585.745	544	2.915		
	Total	1641.855	545			
M_#_MAST	Between Groups	865.304	1	865.304	141.483	.000
	Within Groups	3327.079	544	6.116		
	Total	4192.383	545			

This year marks an anomaly from the marked improvement trend for the treatment school's writing performance. Although still strong, the treatment school no longer enjoys a statistically higher level than that of the control. The associated probability that this event could be due to chance however, is quite high ($p \leq .565$). And, as will be shown in the final 2002 analysis, there is much evidence to indicate that this is a one time random fluctuation. There is very little probability however, that the differences observed in the areas of reading and mathematics can be ascribed to chance ($p<.0005$ in each case).

2001 Means Plots







2002 ANOVA of Core TAAS Objectives

Descriptives

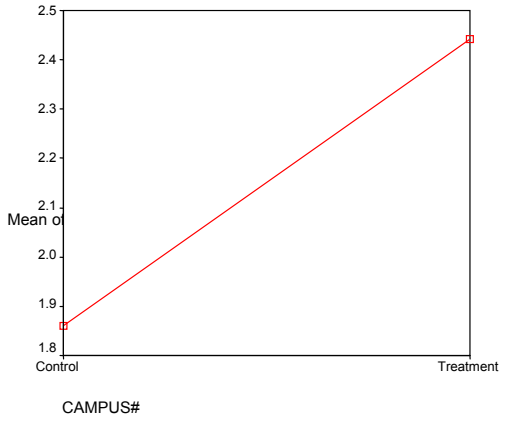
		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval fo Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
W_#MAST	Control	165	1.86	1.035	.081	1.70	2.02	0	3
	Treatment	145	2.44	1.013	.084	2.28	2.61	0	3
	Total	310	2.13	1.064	.060	2.01	2.25	0	3
R_#_MAST	Control	311	4.19	1.929	.109	3.97	4.40	0	6
	Treatment	295	5.33	1.493	.087	5.16	5.50	0	6
	Total	606	4.74	1.822	.074	4.60	4.89	0	6
M_#_MAST	Control	311	7.338	3.3115	.1878	6.968	7.707	.0	12
	Treatment	295	9.776	2.5986	.1513	9.479	10.074	.0	12
	Total	606	8.525	3.2231	.1309	8.268	8.782	.0	12

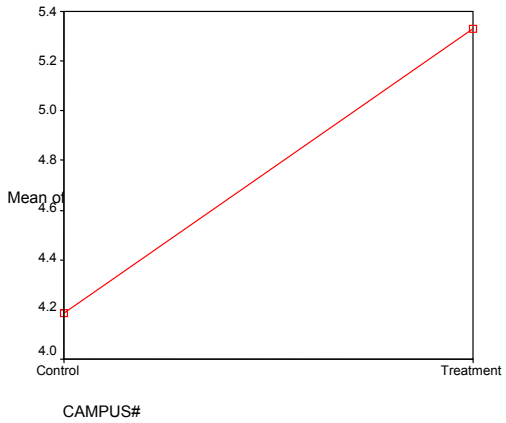
ANOVA

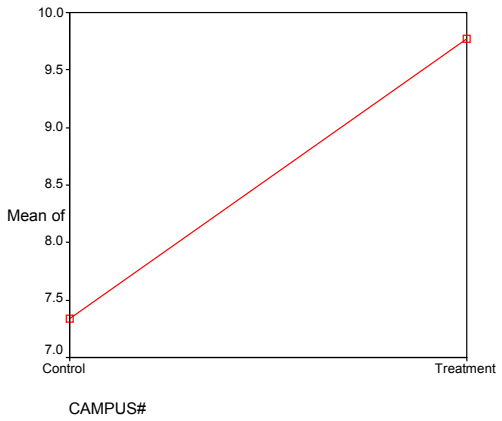
		Sum of Squares	df	Mean Square	F	Sig.
W_#MAST	Between Groups	26.032	1	26.032	24.781	.000
	Within Groups	323.546	308	1.050		
	Total	349.577	309			
R_#_MAST	Between Groups	198.727	1	198.727	66.366	.000
	Within Groups	1808.627	604	2.994		
	Total	2007.355	605			
M_#_MAST	Between Groups	900.345	1	900.345	100.990	.000
	Within Groups	5384.784	604	8.915		
	Total	6285.129	605			

By 2002 the superior student performance at the treatment school is well established and has again achieved statistical significance in all three areas of writing, reading and mathematics. Furthermore, these results again go far beyond the levels typically associated with educational significance ($p \leq .05$) and are each significant at a $p \leq .0005$ level.

2002 Means Plots







The impression left after comparing each of these yearly analyses is of a highly successful program which offers much more then a one-time spike in achievement. Rather, we see a continuous and consistent framework within which student growth and

performance is enabled. It should be noted that due to the high-stakes nature of testing within Texas, the control school was under the same pressure to improve scores as the treatment school. Although we do not know what measures were adopted at the control campus it is clear that the measures they used were demonstrably less effective than the Curry/Samara Model used by the treatment school.

These analyses provided clear empirical, replicable and quantitative evidence that adoption and use of the Curry/Samara Model does indeed provide for a positive impact upon student classroom performance. Furthermore, this positive impact was generated in a school which initially was performing at a lower level than its comparison school - and this comparison school was under the same incentive to improve their own scores due to the high-stakes system of testing and state-mandated accountability for improvement within Texas.

Student-Problem Time Series Analysis

Findings and Discussion

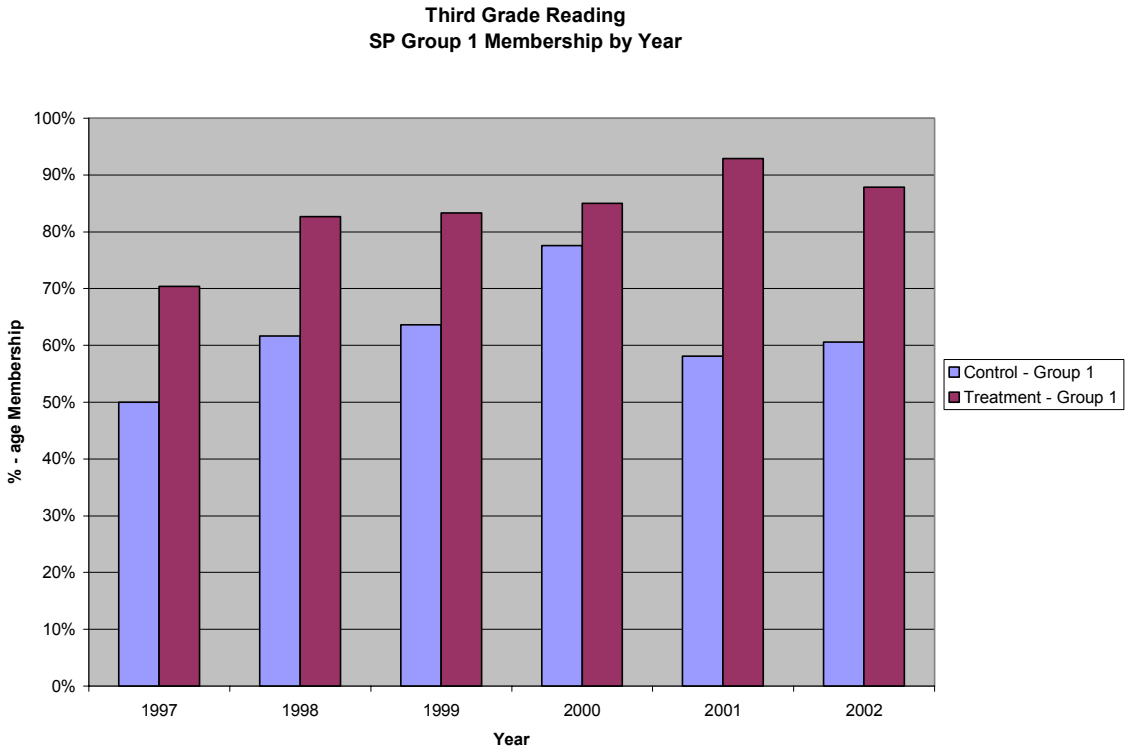
One of the major uses of the Student-Problem Chart lies in its ability to differentiate between students who merely have high scores and those students who achieve such high scores in a consistent fashion. In the S-P Chart high-scoring students are classified on the basis of the item responses into two groups: (Group 1) high-scoring and consistent, these students typically have a firm grasp of the content and are able to make use of their knowledge in a flexible fashion; and, (Group 2) high-scoring and inconsistent, these students have been characterized as having memorized their way

through the content - their understandings are only partial and typically have large conceptual holes.

Similarly, low scoring students are classified on the basis of their item responses into two groups: (Group 3) low-scoring and consistent, these students typically lack substantive grasp of the content and are unable to recall, let alone make use of, their knowledge in any fashion; and, (Group 4) low-scoring and inconsistent, these students are characterized as having unexpected sparks of understanding and occasionally brilliance - these students in many cases have built their own understandings despite ineffective instruction. Occasionally one can identify gifted students from this group.

With this background let us look at what types of students are found at each of the schools. The following discussion will examine the content areas of reading and mathematics in terms of the percentage of students classified in each of these four student-problem groups, in each of the two schools under investigation.

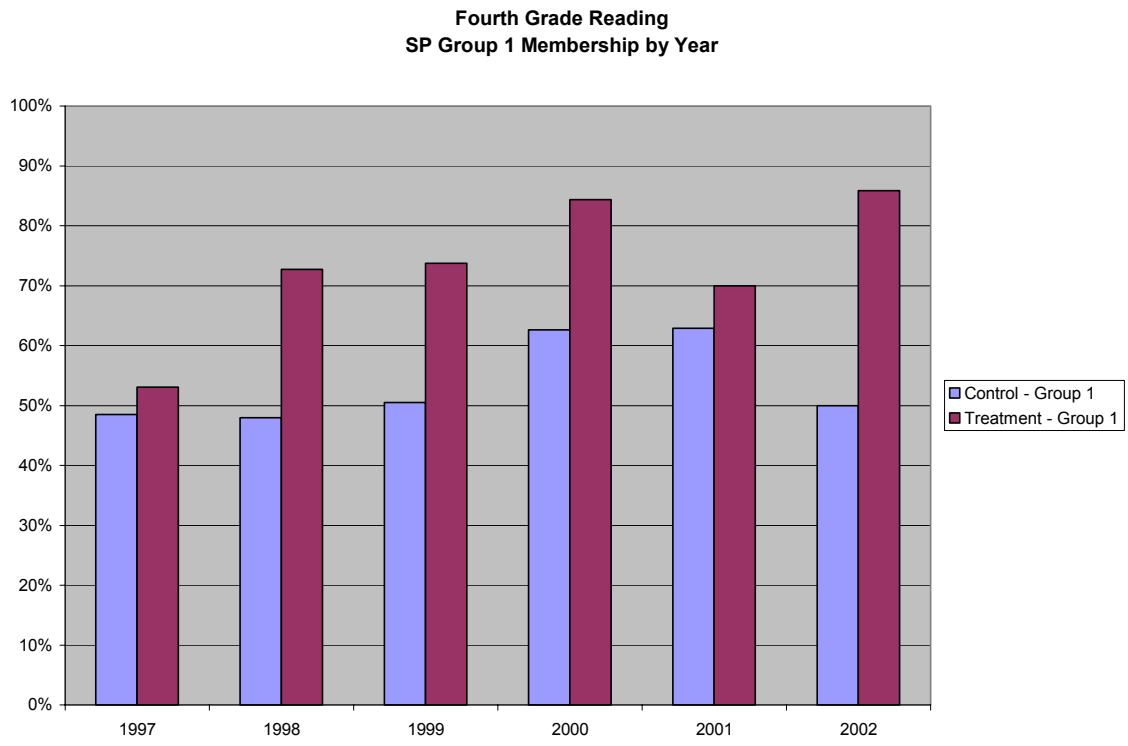
Group 1



This chart provides clear evidence that the differences seen between the treatment and control groups are both statistically and qualitatively significant. Group 1 students are characterized as having a firmer, more flexible understanding of academic content. To see that students from the treatment school are not just scoring higher, but are being placed into S/P Group 1, is a strong indicator that the intervention does more than merely raises scores, it also raises understanding and higher order thinking.

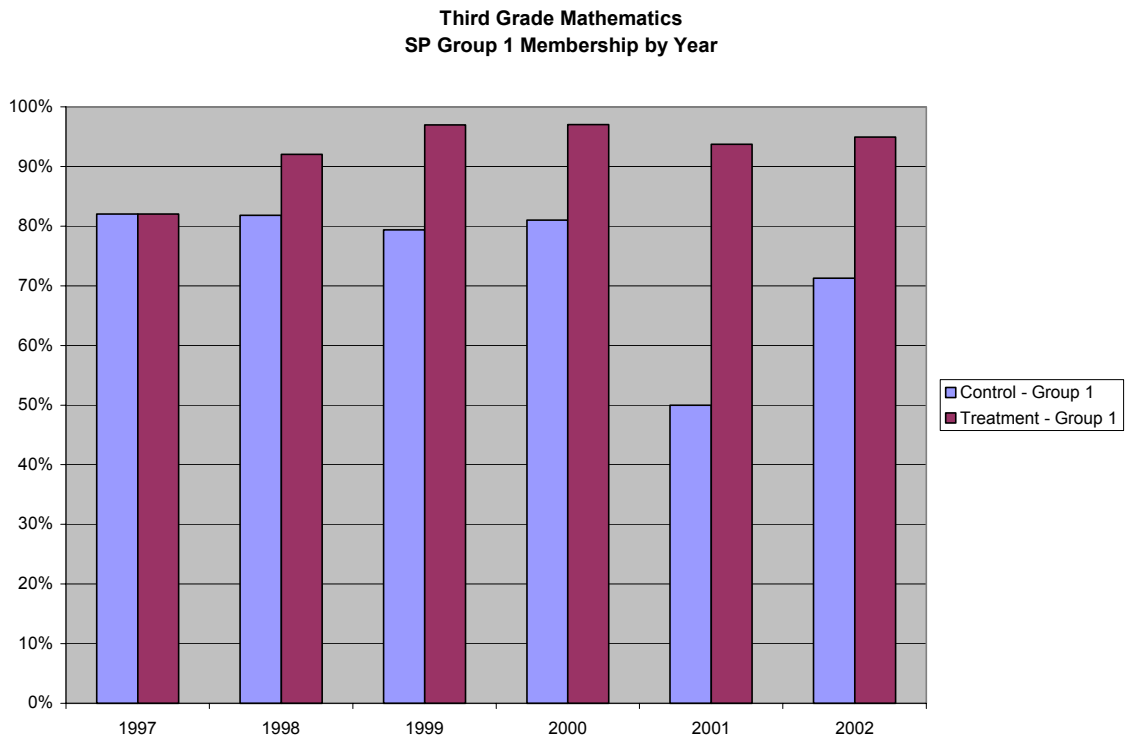
The remaining charts for the cases of fourth-grade reading, third-grade mathematics, and fourth-grade mathematics each provide additional evidence of this consistent growth across grade levels and content areas[4].

[4] Due to the low number of objectives in the writing domain (3) it is not possible to construct an SP Chart for this area. An SP Chart requires sufficient numbers of both students and objectives for statistical differences to be ascertained.

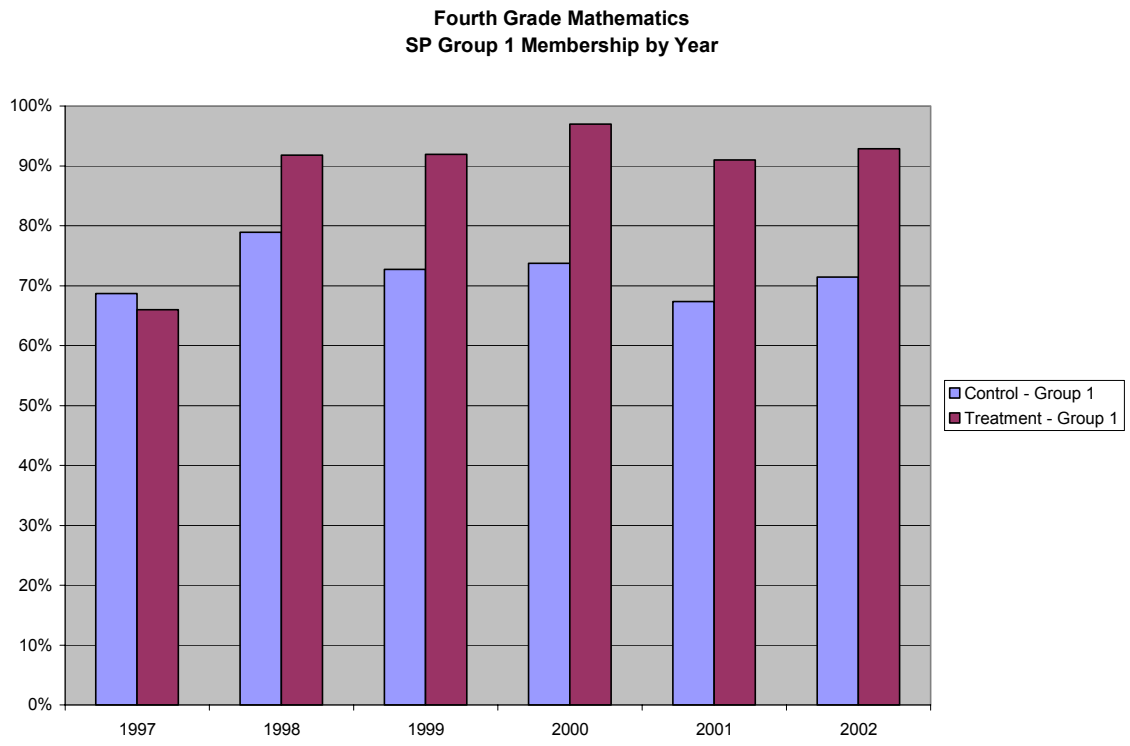


This chart provides an interesting insight into the nature of growth shown in both schools. The treatment school shows an increase in students classified in Group 1, while beginning with the year 2000 we see the control school showing a marked decrease in Group 1 students.

Typically when seeing a decrease of this type in Group 1 students it is due to an increased reliance upon drill, practice and memorization as primary teaching strategies. Such practices are common in schools which focus on test taking strategies as opposed to adequately covering the content. Unfortunately, such drill on test taking strategies was highly effective under the TAAS test and as such was adopted by many schools at the expense of offering a rich curriculum. This strategy is unlikely to be successful under the new state mandated Texas Assessment of Knowledge and Skills (TAKS).

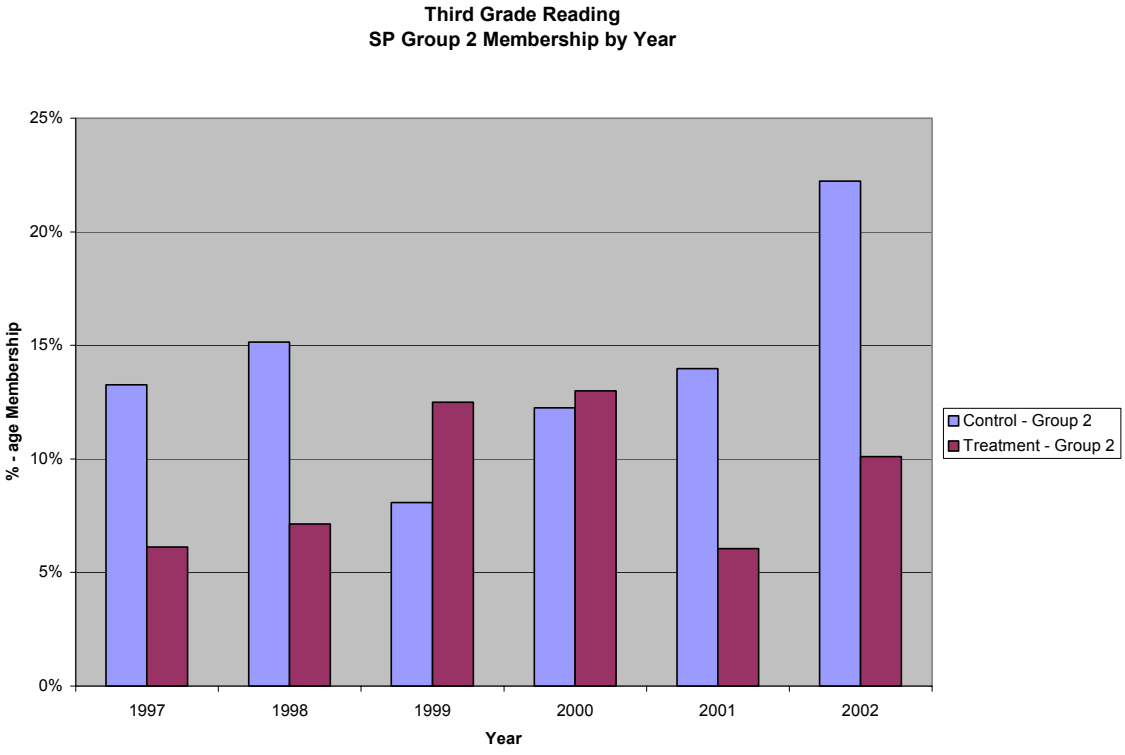


In this chart we see the same superior performance from the treatment school. There is an interesting decline from the year 2000 to 2001 however, for the control school students. A probable interpretation in this case is a concern over lower test scores in the year 2000. Concerns such as this often leads schools lacking a strong curriculum commitment, to adopt a didactic and memory driven instructional sequence coupled with the earlier mentioned emphasis upon test-taking strategies. The subsequent increase from year 2001 to 2002 could either indicate a successful matching with portions of the test or dropping the didactic and memory driven instructional sequence. Irregardless, however, memory driven instruction is likely to be unsuccessful in the new curricular grounded TAKS.



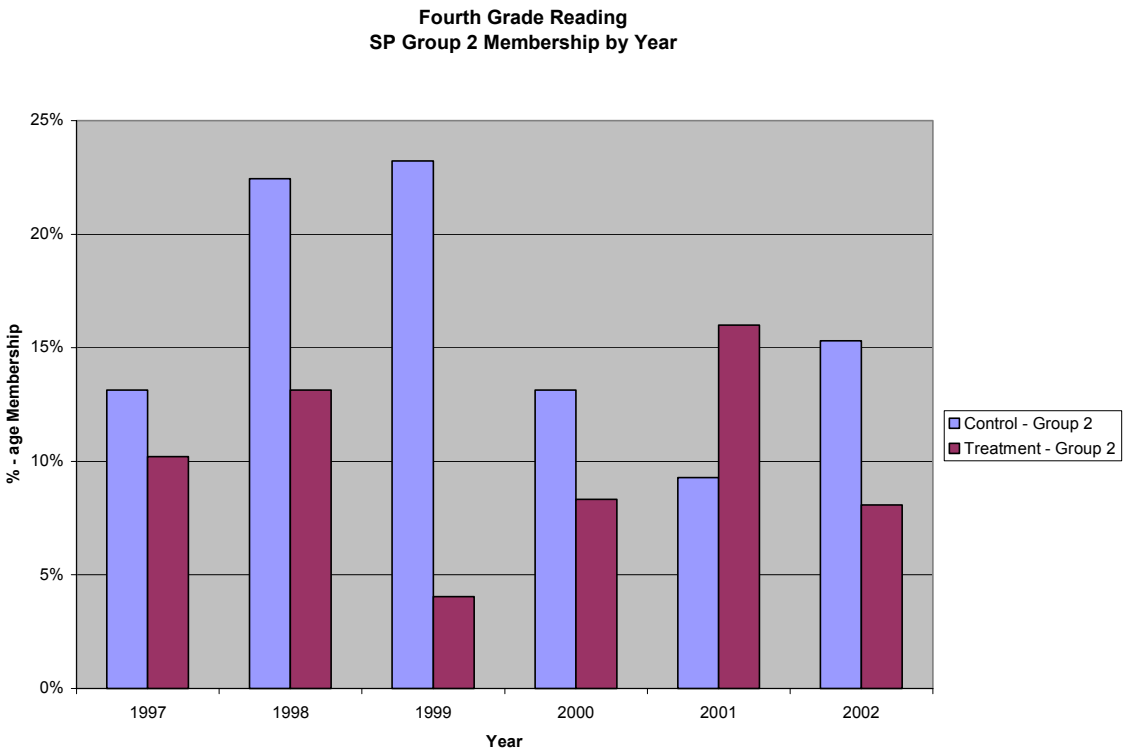
In this chart we again see greater numbers of Group 1 students being generated in the treatment school, while the decline in performance from 2000 to 2001 is once more observed for the control school's students. The presence of this drop at more than one grade level lends credence to the notion that the school adopted a more memory driven instructional system.

Group 2

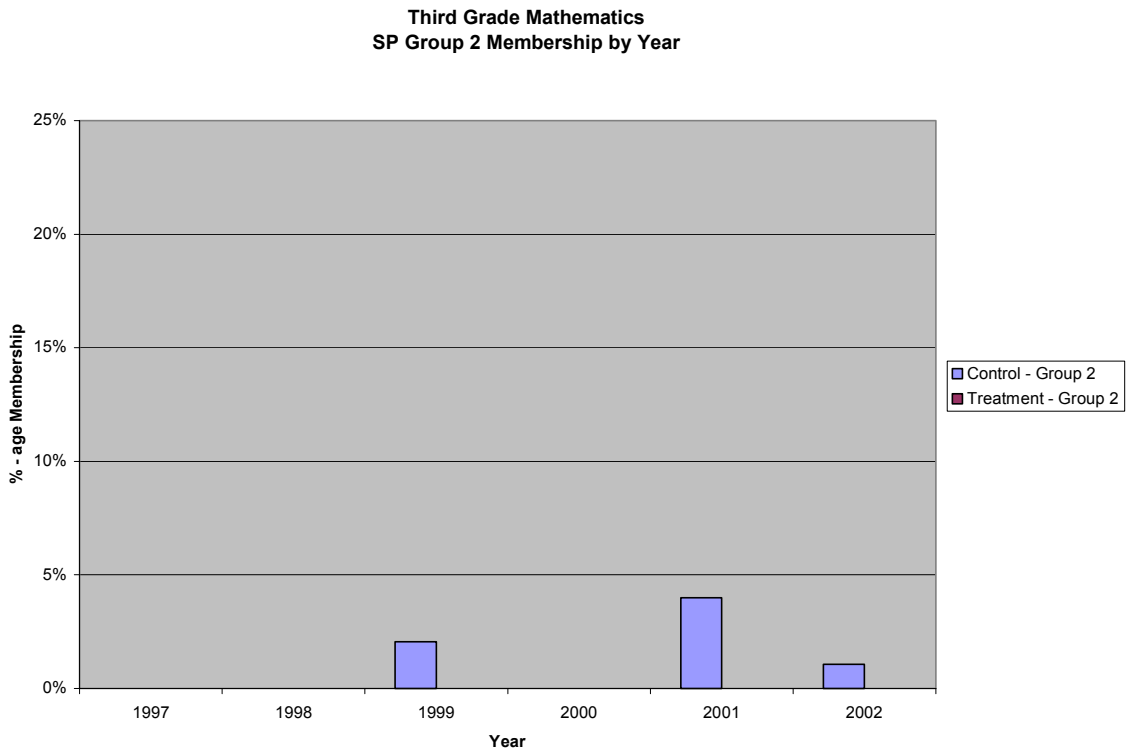


A goal of successful instruction, from the earlier discussion of S-P Group characteristics, is to attain high numbers of students in Group 1 and low numbers of students in Group 2. In this chart, for the case of the control school, the exact opposite has occurred. In this school the number of students classified in Group 2 increases each year following 1999. Given that Group 2 is categorized as having relied on memory much too often, and for much too long, the picture of instruction which emerges is that of a system driven by memorization of isolated facts.

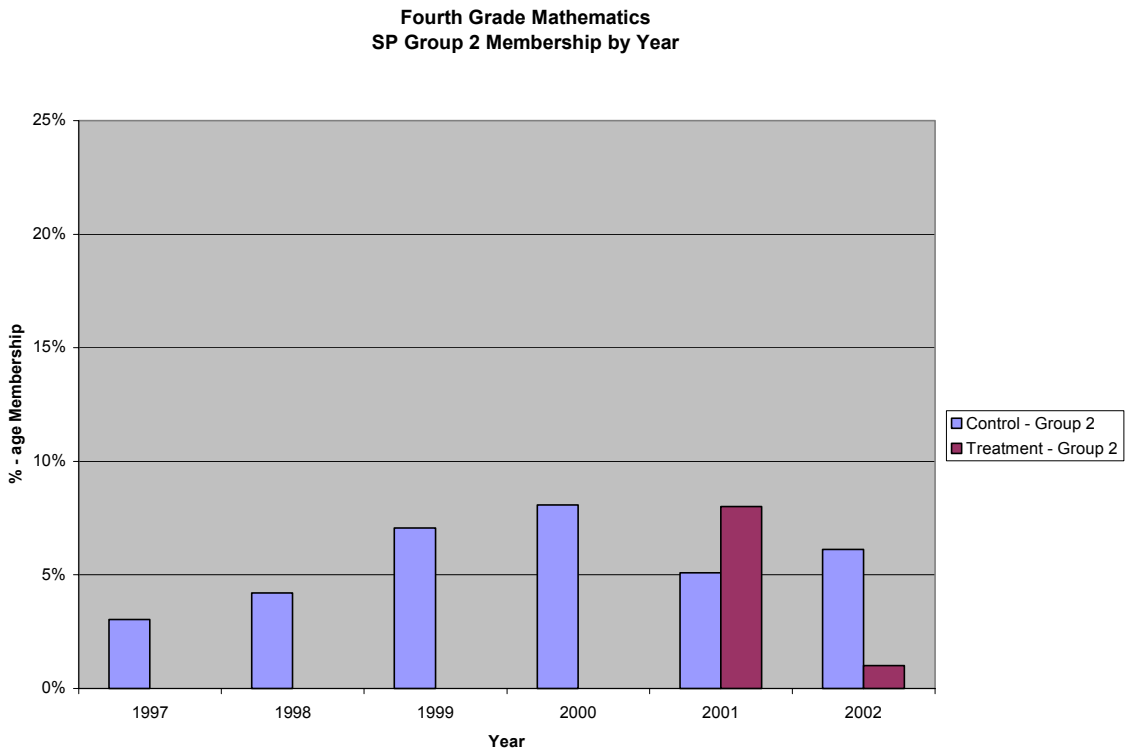
This picture is supported by each of the following charts across both grade level and content area.



It would be interesting at this point to interview teachers at both campuses to see what happened in the Reading Curriculum during the years 2000 and 2001. Based upon the S/P chart we see an unhealthy rise in Group 2 students for the treatment school. This increase, however, appears to have turned around during the 2001-2002 school year. In examining the data it seems likely that this rise is due solely to the fourth-grade team. We see a similar rise in Group 2 students in the mathematics area of fourth-grade as well. This could indicate a need to focus upon mathematics with this fourth-grade team.



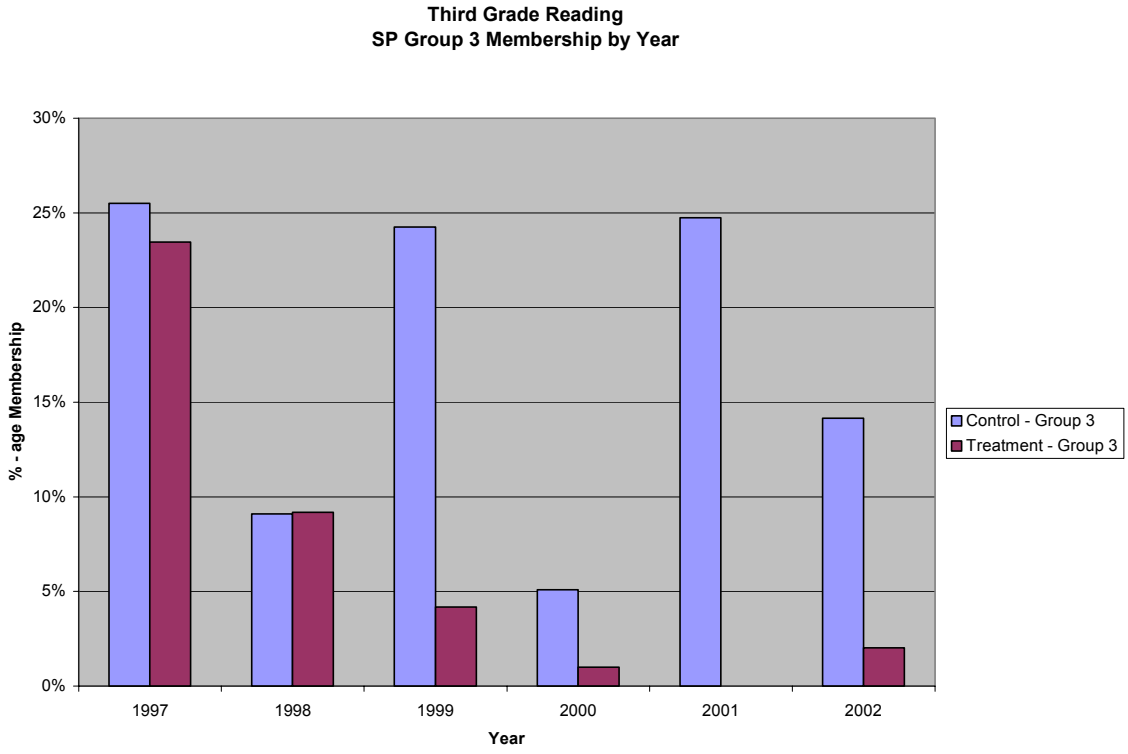
This chart is fairly self-explanatory. It should be noted that much of the third-grade mathematics curriculum deals with operations which are learned via memorization. If one assumes the typical memory driven mathematics curriculum, it is not too large of a surprise that the instructional system provides reasonable results for the objectives as they are measured by the TAAS test.



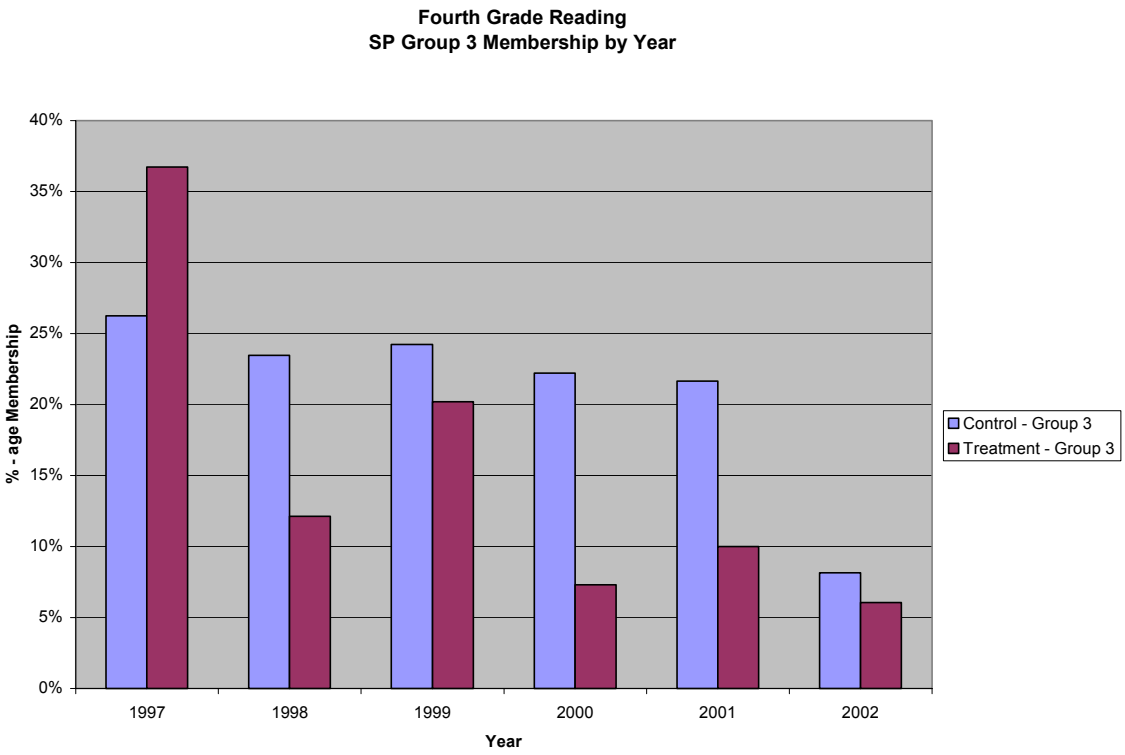
The larger numbers of students in Group 4 shown in this chart indicates in many ways the increasing complexity of the mathematics curriculum and the sophistication with which it was measured in the fourth-grade TAAS test.

Group 3

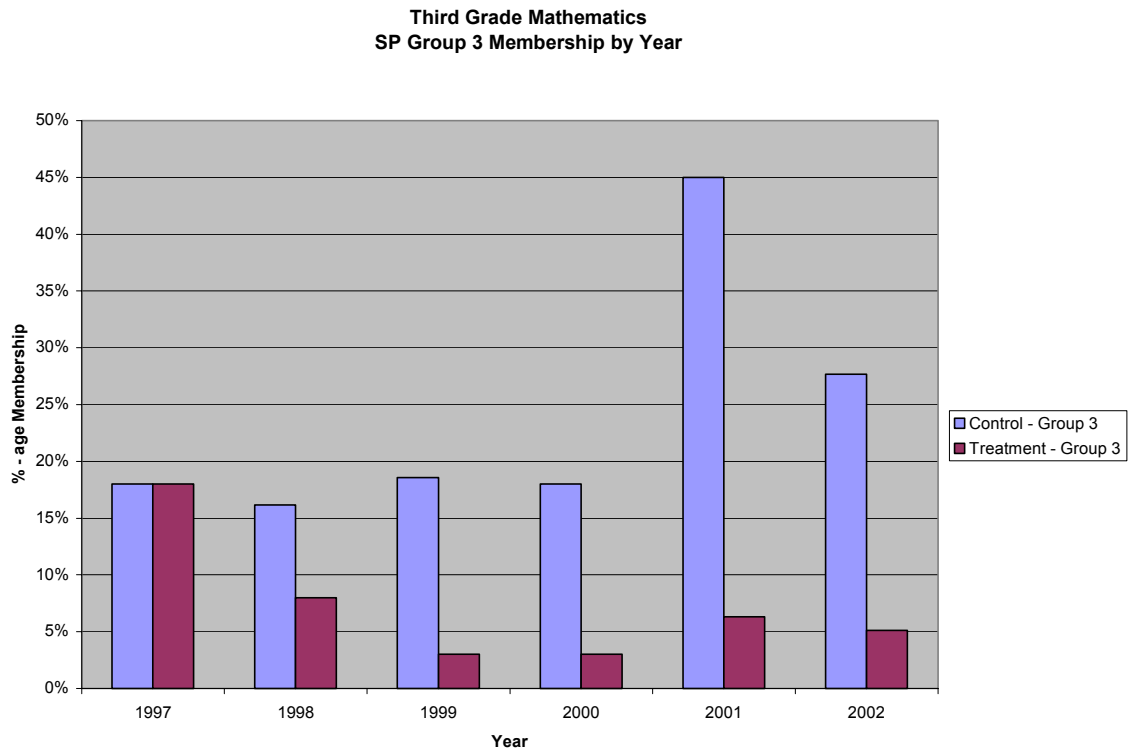
Given that Group 3 students are characterized as having neither procedural understanding, nor the ability to recall or use their knowledge in any meaningful fashion, a goal would be to have the smallest number of students possible in this group. As such the following charts are fairly self-explanatory.



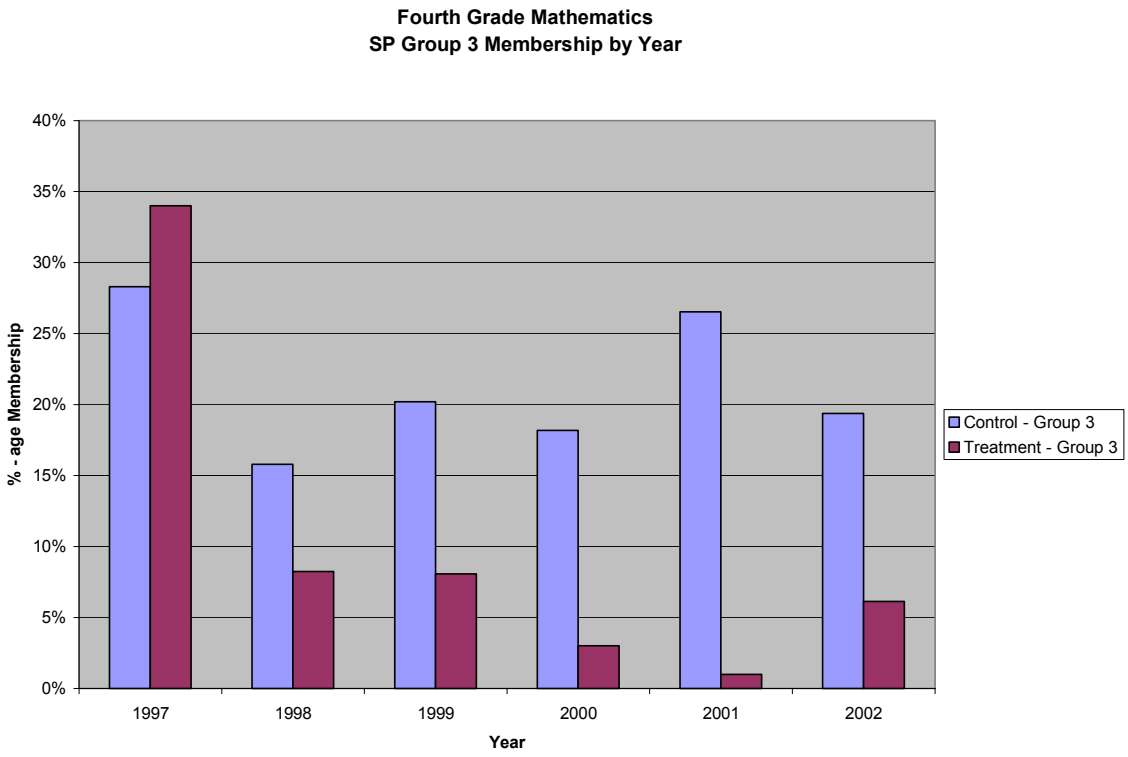
The near absence of Group 3 students in the Curry/Samara school is striking, particularly given the 15 to 25 percent levels present in the comparison school. This speaks to a program which is effective in reaching low-performing students as well as meeting the needs of the high-performing (as was evidenced in the Group 1 analysis provided earlier).



The constant decline of Group 3 students in the treatment school is a positive sign, particularly given the nearly constant decreases shown each year. Evidence of a shift from Group 2 to Group 1 as a result of classroom experience using the Curry/Samara Model may also be observed.



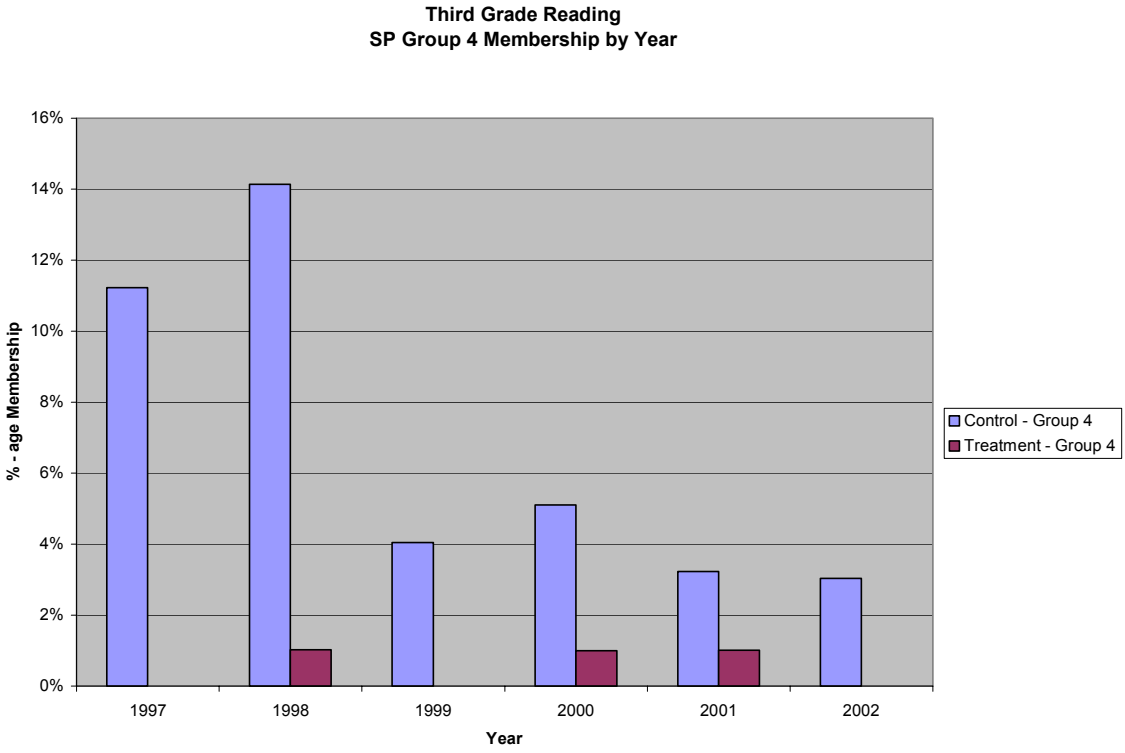
The marked increase in the year 2001 in Group 3 at the comparison school serves as yet another indication of a failed reliance on a memory-recall teaching strategy. The continuing decline and relatively low numbers in the Curry/Samara school seem to indicate a more holistic and healthy curricular emphasis.



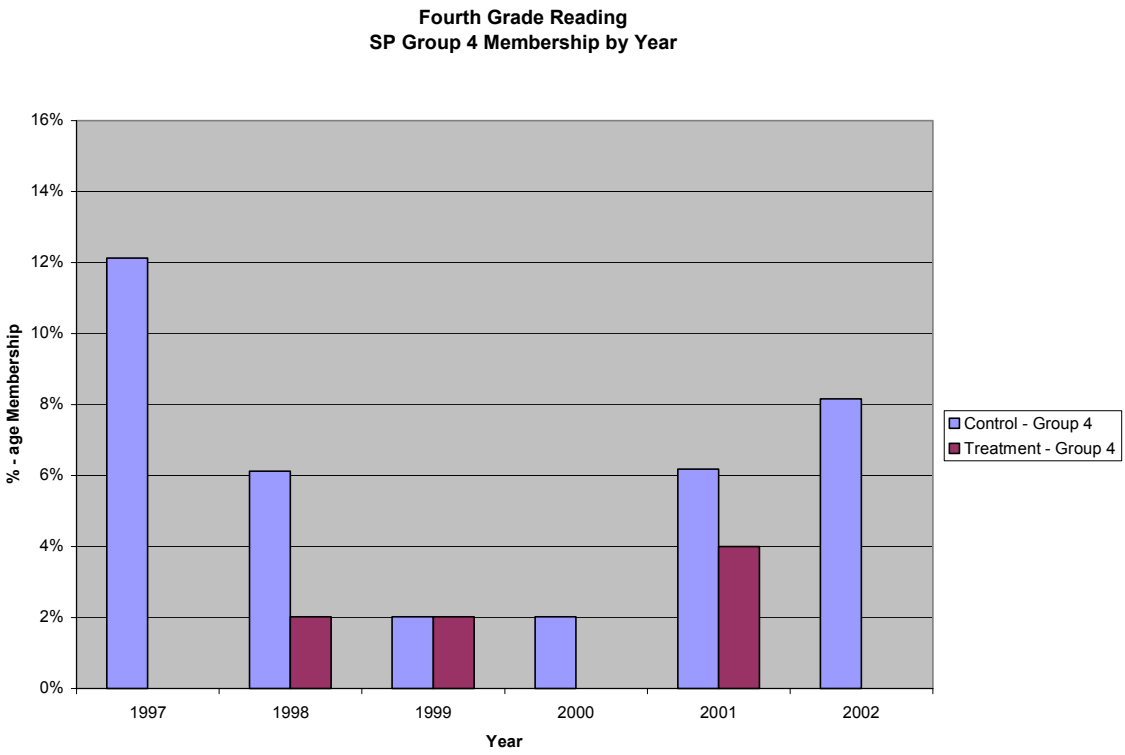
This picture in mathematics is repeated again in Grade 4.

Group 4

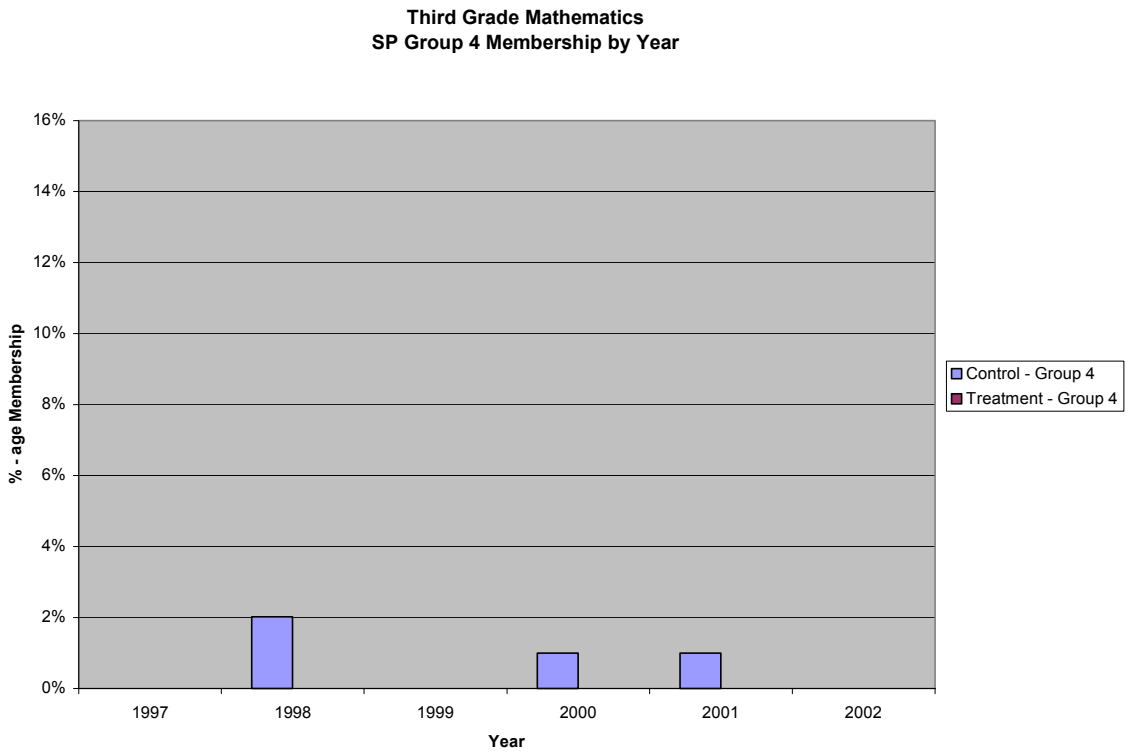
In interpreting charts for S/P Group 4 it is helpful to remember that these students are often those who must make sense of the content on their own. This is due either to a markedly different manner of understanding than that expected in the classroom, or the absence of effective instruction. Although the data does not provide sufficient information to distinguish between these two events, the evidence thus far tends to indicate increasing reliance upon a memory driven curriculum and instruction system which is not meeting the needs of the students. One result of this would be an increase in the number of Group 4 students.



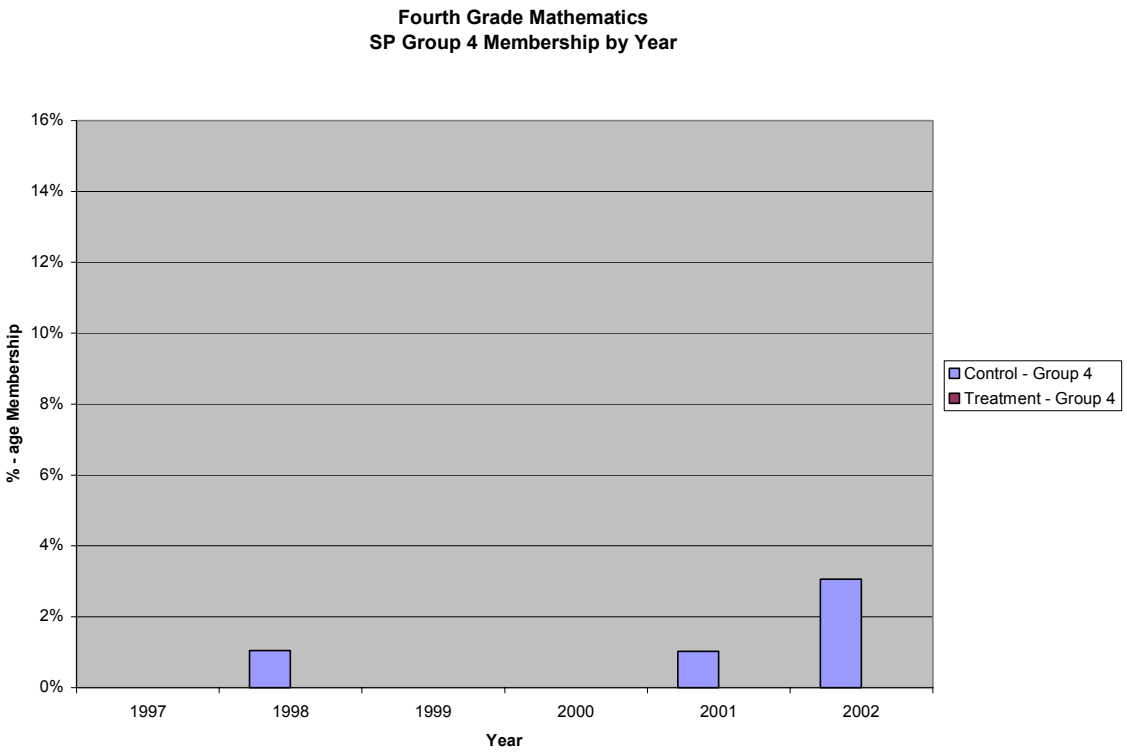
The initially high levels of Group 4 students at the beginning of this time series are quite surprising. It would be interesting to do a follow-up study to determine precisely what was occurring in the curriculum in the years 1997 and 1998 at the control school. The picture which emerges is an instructional sequence, which while on the whole ineffective, was successful in generating partial and incomplete understanding in some key areas.



The same instructional sequence, successful in generating partial and incomplete understanding in some key areas, appears to be in effect at the control school in the fourth-grade as well during the years covered by this analysis.



The total absence of Group 4 students in the treatment school implies that partial understandings appear to not be generated using this method. It appears that full and complete sets of understandings were developed.



This picture is repeated again at the fourth-grade level, lending further credence to the successful of the Curry/Samara Model in generating full-blown, as opposed to partial, understandings.

An alternative series of representations showing this same data is presented in Appendix B, which is available upon request (call 800.867.9067 for more information). In these charts all four S/P groups are presented by school over the six-year intervention. These charts are more powerful, but require more processing to fully understand. It would be worthwhile for the reader to examine them at this point to see how well they serve to convey the same story.

Student-Problem Growth Curve Analysis

Findings and Discussion

There are two lines that can be drawn on the S-P Chart. The S-curve is drawn by placing a solid vertical line corresponding to the total test score earned by the student represented by that row. The S-curve is completed by starting at the bottom of the chart and connecting the top end point of each vertical line segment to the bottom end point of the line segment to the left end point of the line segment above it.

The S-curve provides a visual display of the proficiency level of the students in the classroom. A vertical S-curve represents a homogeneous classroom (one in which most of the students are performing very similarly), where a diagonal S-curve is evidence of a heterogeneous classroom (a wide range of performance). An S-curve shifted to the right shows a high proficiency level in the classroom.

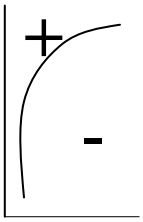
The P-curve is drawn in a similar manner except that the roles of students and items are reversed. A horizontal line segment is drawn which corresponds to the number of students correctly answering the item represented by that column. Starting at the bottom of the chart and connecting the right end point of the line segment above it completes the P-curve.

In an ideal classroom with an ideal test (where everything tested is taught and learned) the S and P-curves would coincide. In actual practice, however, it is quite common to have the S-curves and P-curves diverge a little from each other, due to individual differences commonly observed in the classroom. For example, if the test is measuring information not covered in the classroom the divergence between these two

curves will be increased. A large divergence between the S and P-curves signals a possible mismatch of test objectives and instructional objectives in the classroom.

With this background it is possible to imagine a multitude of uses for the SP Chart. One such use, which has shown to be very helpful in field based projects, is to document classroom improvement over time. Often in the initial stages of an intervention there has not been sufficient time for significant gains to be shown on standardized testing measures, yet both the researcher and the classroom teacher are in complete agreement that tremendous progress is being made. The dilemma becomes one of documenting this growth in a form that is communicable to others.

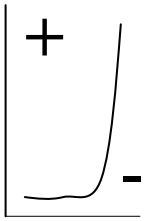
At this point the SP Chart can provide a tremendous insight into the classroom changes underway. For simplicity, this example will look only at the S-Line as defined in the preceding sections. As we examine S curves at the beginning of an intervention patterns similar to the following are often observed:



In examining this curve a few observations may be made. First, the relative number of items mastered by these students is relatively small and the number of students successfully mastering items is likewise small. Secondly, in order for this class to show real growth an increase in the number of mastered items must be achieved, as well as a marked increase in the number of students showing mastery.

This is in marked contrast to many administrative foci that look for simple gains in scores without concurrent conceptual mastery. For example, it is entirely possible to imagine two S-Lines drawn from two different classrooms, identical in their overall shape to that shown above. Imagine now that the mean scores in these two classrooms differ by a statistically significant amount. A simplistic interpretation would be that the higher score, as reflected by the mean, would be indicative of the better classroom. However, if all that is demonstrated is a higher score, and the underlying S-curve remains the same, then one would be hard pressed to argue that real growth has been achieved. In this case we began with a small group of highly able students, who despite their instructional history, have managed to master the content. The vast majority of the class has not. We may see an increase in the mean classroom score, yet still have the same relatively small numbers of students mastering the content.

Let us now consider a second curve typical of many intervention projects.

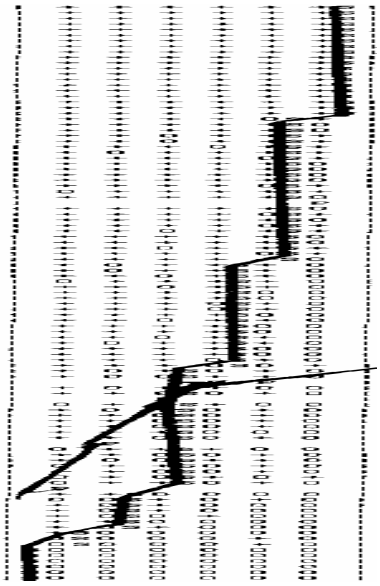


In examining the S-Curve we can easily see that despite a lack of mean increase, we see a large number of students who are progressing in their mastery of concepts. This classroom is rapidly becoming an environment within which meaningful instructional dialogue may take place. For the cases shown in this report we see growth along the lines

of this curve type, coupled with statistically significant improvements in student performance – truly an ideal case.

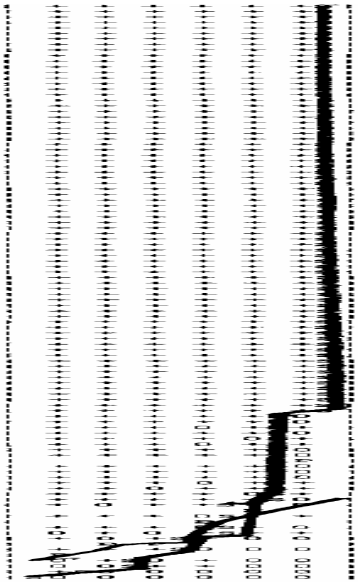
Taking these two different curve types as models there is a great deal of evidence concerning the nature and type of growth provided within the Curry/Samara Model from the S-P Chart. Contained in Appendix C (available upon request, call 800.867.9067) are illustrations created from the SP Charts used in this phase of the analysis. These were created by reducing the font size of the original response matrix until it fit on one page, then drawing in the S and P curves. This was then scanned, copied into Word and resized to a standard 3" x 2" image to standardize the graphic and make differences easier to observe and document. The result of this process was a powerful set of images to use in describing changes in S-curves and P-curves in a program showing increasing growth (the treatment school) relative to one staying static (the control school).

This data provides multiple opportunities to provide explicit examples of the type of growth curves hinted at in the above discussion. For example consider the curve constructed from the S-P chart analysis of *2002 TAAS Reading* for the control school's Fourth Grade.



Note the nearly perfect diagonal of the student response curve; this is much more in keeping with an instructional system where students are learning at random rather than by an effective instructional strategy. Furthermore, the large divergence between the S-curve and P-curve is a classic example of the ineffectiveness of this instruction in meeting the needs of the students and the demands of the curriculum. Such a large divergence between the S and P-curves signals a possible mismatch of test objectives and instructional objectives in the classroom**.**

To see a more positive picture, such as observing a large number of students who are progressing in their mastery of concepts and creating an environment within which meaningful instructional dialogue may take place, one need only look to the multiple examples which may be drawn from the treatment school. One such example is the curve constructed from the S-P Chart analysis of *2002 TAAS Reading* for the treatment school's Fourth Grade – the same grade level and year as that used in the previous illustration.



This curve illustrates not only a very high degree of success, but that the curriculum is aligned closely with the evaluation.

Conclusion

Whether examining the ANOVA analyses, the preliminary S-P group analysis, or the Student-Problem Curve analysis the end implication is clear. We see a highly successful program which offers a continuous and consistent framework within which student growth and performance is enabled. Furthermore, this program creates statistically significant and measurable results when compared with a comparable population. Nor are these results merely of statistical importance. Based upon the Student-Problem Chart numerous qualitative differences were found regarding the nature and type of student learning as well as the instruction being offered in these two systems.

In each case the Curry/Samara Model provided for a significantly richer environment within which learning takes place.

In interpreting these findings it must be recognized that there exists an extremely high-stakes testing environment within the state of Texas. Schools are praised or shut down in disgrace based upon their performance on the TAAS test from which this data was derived. Both of the schools in this analysis were under the same pressures to improve scores. Although we do not know what measures were adopted at the control school to improve student achievement, it is clear that the measures they used were demonstrably less effective than the consistent gains offered by the Curry/Samara Model, as implemented at the treatment school.

The results of these analyses provide clear empirical, replicable and quantitative evidence that adoption and use of the Curry/Samara Model does indeed provide a positive impact upon student classroom performance. Furthermore, this positive impact was generated in a school which initially was performing at a lower level than its comparison school - and this comparison school was under the same incentive to improve their own scores due to the high-stakes nature of testing and state-mandated accountability for improvement within the state of Texas.

Appendices

A) Demographics of schools used within the report.

B) Time Series SP Analysis in Chart Form.

C) Student-Problem Growth Curve Analysis.

Appendix B & C are available on request, please call 800.867.9067 for more information.

Appendix A

Demographics of schools used within the report.

GENDER * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME		Total
			Control	Treatment	
3.0	GENDER	F	52	65	117
		M	63	64	127
	Total		115	129	244
4.0	GENDER				
		F	53	63	116
		M	56	51	108
	Total		110	114	224

TITLE1A * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME		Total
			Control	Treatment	
3.0	TITLE1A	not identified economically disadvantaged	4	1	5
		Student attends campus with schoolwide program	110	128	238
		Student participates in program at assistance school	1	0	1
	Total		115	129	244
4.0	TITLE1A	not identified economically disadvantaged	3	0	3
		Student attends campus with schoolwide program	106	114	220
	Total		109	114	223

MIGRANT * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME: Control	CAMPNAME: Treatment	Total
3.0	MIGRANT	0	115	129	244
	Total		115	129	244
4.0	MIGRANT	0	109	114	223
	Total		109	114	223

LEP * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME: Control	CAMPNAME: Treatment	Total
3.0	LEP	Not Limited English Proficient	100	95	195
		Limited English Proficient	15	34	49
	Total		115	129	244
4.0	LEP	Not Limited English Proficient	102	66	168
		Limited English Proficient	7	48	55
	Total		109	114	223

BI_LING * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME: Control	CAMPNAME: Treatment	Total
3.0	BI_LING	Not Participating in Bilingual Program	115	118	233
		Participating in Bilingual Program	0	11	11
	Total		115	129	244
4.0	BI_LING	Not Participating in Bilingual Program	109	112	221
		Participating in Bilingual Program	0	2	2
	Total		109	114	223

RACE * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME		Total
			Control	Treatment	
3.0	RACE	Asian or Pacific Islander	1	2	3
		African American	48	9	57
		Hispanic	24	95	119
		White, not of Hispanic Origin	42	23	65
	Total		115	129	244
4.0	RACE	Asian or Pacific Islander	1	0	1
		African American	46	3	49
		Hispanic	24	94	118
		White, not of Hispanic Origin	38	17	55
	Total		109	114	223

AT_RISK * CAMPNAME * GRADE Crosstabulation

Count

GRADE			CAMPNAME		Total
			Control	Treatment	
3.0	AT_RISK	Not Designated At Risk of Dropping Out	85	46	131
		Designated At Risk of Dropping Out	30	83	113
	Total		115	129	244
4.0	AT_RISK	Not Designated At Risk of Dropping Out	64	40	104
		Designated At Risk of Dropping Out	45	74	119
	Total		109	114	223