

Partners in Character Education Program

Evaluation Report
2006 – 2008



Floresville, Texas

Partners in Character Education Program
Evaluation Report 2006 to 2008

September 15, 2008

Superintendent

David Vinson

Project Director

Leanne Marshall

Grant Evaluator

Scott Walker

Floresville Independent School District
908 10th Street
Floresville, Texas 78114
(830) 393-5300



Floresville Independent School District (FISD) is a rural district consisting of six campuses that accommodates approximately 3,700 students. In 2007, 253 students graduated from Floresville High School and the District received an Acceptable Rating from the Texas Education Agency. This is a remarkable achievement considering the diversity and socio-economic status of our students. According to the 2000 Census Bureau, of the population of Floresville citizens who are 25 years and older, only forty-three percent graduated from high school and the per capita income is only \$13,340. Floresville ISD is not a wealthy district.

In the spring of 2005, a student survey, The Ethics of American Youth, was administered to FISD students and after evaluating these results, inconsistency between students' words and actions regarding character was evident. A Character Education Committee was formed to address this concern and design a Character Education Program for our school. A task resulting from this committee was to seek external assistance to aid school teachers and administrators in tackling this problem.

The FISD Character Education Program today involves a number of initiatives targeted toward students, teachers, school administrators, parents, and the community as a whole. "Character" is on the forefront of our minds as we progress through our day. "Character" is embedded in our lessons, is actively addressed by our community, and is now considered a moniker of our district, largely due to the funds provided by this grant program and the dedicated work of FISD's staff, faculty, and our community.

David Vinson, PhD
Superintendent

1. BACKGROUND AND PURPOSE

1.1 DESCRIPTION OF THE CHARACTER EDUCATION PROGRAM

The Floresville Independent School District's (FISD) Character Education Program has taken a multi-faceted approach to character education. A core group of 24 highly-motivated "lead" teachers have taken on the development of new "character education" curricula. Each of these teachers has developed a subject-specific lesson set with a character education component (see <http://www.floresville.isd.tenet.edu/CharacterED/>). These teachers and their dynamic curricula have served as models and mentors for other FISD teachers.

Likewise, a district-wide character education professional development program was established to introduce teachers at all levels to character education. A school administrator character education professional development program also was launched to thoroughly cover aspects and challenges of character education in a public school setting. Thus, character education has become a flagship theme in FISD.

An expanded service learning program was established that includes middle school students rather than just high school students. And, a middle school-to-high school transition program was established to bridge the critical gap students must address when they enter high school.

In addition to providing teacher, staff, and administrator character education professional development and expanding student programs related to character, the community has been offered parenting workshops and the opportunity to become involved in student character through involvement in FISD's long-standing Character Committee. The committee is made up of parents, teachers, administrators, and local community members.

FISD has taken on the complex task of implementing these general categories of programming that are measured in terms of their program goal and three objectives, each with its own sub-objectives.

1.2 GOAL AND OBJECTIVES

Floresville ISD (FISD) has had a long-standing, district-wide goal to address student character through the classroom and other school-related events. Summarized, the program goal is to "Build a Community of Character." Three objectives were established to achieve this goal:

Objective 1: Floresville ISD will integrate instruction in the character traits of: (1) trustworthiness, (2) respect, (3) responsibility, (4) fairness, (5) caring, and (6) citizenship into our state-mandated academic curriculum.

Objective 2: Floresville ISD students will contribute to the community and enhance character traits through service learning projects.

Objective 3: Floresville ISD will serve as a Character Education model by providing our youth with a caring environment that supports the integration of excellence and ethics in all phases of life.

1.3 FACTORS SELECTED FOR EVALUATION

The program's evaluation plan was established by means of identifying strategies designed to achieve each objective. Each strategy has a process evaluation component and an outcome evaluation component. The methods by which these two evaluation components are conducted include records analyses, content analyses, randomized focus groups, interrupted time-series and single-group time series investigations, as well as quasi-experimental and randomized experimental methods. The measures and instruments themselves thus conform to the evaluation approach for each strategy to include original instruments that have been highlighted in the program's successes in Section 1.4.

1.4 PROGRAM SUCCESS

Of overall noteworthiness is the fact that Floresville Independent School District (FISD) students, teachers, staff, and administrators all have a heightened awareness of character education due to the implementation of this grant program. More specifically, the rigorous evaluation measures applied have resulted in solid, valid instruments that can be used in other schools across the country. For instance the Classroom Community Environments Survey (CCES), specifically developed for this program, has yielded statistically valid teacher and student versions. Likewise, the Service Learning Survey (SLS) has consistently proven reliable, as has the High School Transition Survey (HSTS).

Notable results from the measures themselves include statistically significant differences between the baseline administration of the CCES-Teacher and the second administration. Measures using the same instrument also indicated statistically significant differences between the control group of teachers and the experimental group. Further, scores from the Powerful Teaching test of teacher knowledge, awareness, and efficacy continued to climb steadily over three instrument administrations.

The CCES-High School version detected significant effects between November 2007 and February 2008 within the control group. There were also statistically significant differences between service learning experimental group students and their control group counterparts when using the SLS. Students enrolled in service learning strongly indicated that they perceived they had the power to make a difference in their community.

2. METHODS

2.1 STUDY SETTING

This study was conducted in Floresville Independent School District (FISD) in south-central Texas. Floresville’s population is 7,250 and continues to experience growth at 23.6% annually. Floresville is the county seat of Wilson County with a population of 32,408 (U.S. Census, 2000).

FISD is the largest of five school districts in Wilson County, with over 3,600 students, 61% of whom are minority. FISD has a completion rate of 93%, despite 54% of the students listed by the State of Texas as economically disadvantaged and 41% labeled at risk attending schools where the annual teacher turnover rate is 17.4%. Of the 254 teachers in FISD, 80% are minority. Teachers’ average years of experience, despite the high turnover, remains at 11 years. (Texas Education Agency, 2007)

The teacher population is small enough that after the 2006-2007 school year we discovered a “wash out” effect in our in-district control groups. In other words, our control group teachers, and subsequently their classes, could not avoid the character education messages and curriculum being initiated. Therefore, for the 2007-2008 academic year we negotiated an agreement with surrounding school districts (Poth ISD, Stockdale ISD) to act as our control groups.

Table 2.1. Control group school demographics

	Poth ISD	Stockdale ISD
Number of students	242	212
% Minority	41%	47%
Economically disadvantaged	25%	34%
At-Risk	30%	45%
Teachers	23	24
Avg. yrs. experience	13	13

(TEA, 2007)

2.2 DESCRIPTION OF HOW OUTCOMES WERE MEASURED

The Character Education Program consists of three primary objectives and a number of secondary objectives (Table 2.2). Each objective has a question to answer and it follows that each question has a measure, or set of measures, that have been implemented.

Table 2.2. Program objectives, questions, and measures

Objective	Question	Measure
-----------	----------	---------

Objective 1: Floresville ISD will integrate instruction in the character traits of: (1) trustworthiness, (2) respect, (3) responsibility, (4) fairness, (5) caring, and (6) citizenship into our state-mandated academic curriculum.	To what extent did teachers' knowledge and awareness increase?	<i>Classroom Community Environments Survey (CCES)</i>
	To what extent did students internalize the six character trait?	CCES: Elementary Version Middle School Version High School Version
	To what extent did special education students internalize the six character traits?	CCES: Elementary Version Middle School Version High School Version
Objective 2: Floresville ISD students will contribute to the community and enhance character traits through service learning projects.	To what extent... (1) did students learn (2) students' attitudes change (3) were students impacted by service learning?	<i>Service Learning Survey (SLS)</i> Three scales: 1) Learning 2) Attitudes 3) Experience/ Impact
Objective 3: Floresville ISD will serve as a Character Education model by providing our youth with a caring environment that supports the integration of excellence and ethics in all phases of life.	To what extent did teachers' knowledge, awareness, and efficacy increase?	10-Item Test (primary/elementary teachers) a) Knowledge of principles of 40 Dev. Assets b) Awareness of 40 Dev. Assets c) Efficacy in teaching about 40 Dev. Assets.
		22-Item Test (middle school teachers) a) Knowledge of Powerful Teaching b) Awareness of Powerful Teaching c) Efficacy in Powerful Teaching.
	To what extent did participants' knowledge, awareness, and efficacy increase?	<i>Building Developmental Assets Test</i>
	To what extent did participants' knowledge, awareness, and efficacy increase?	<i>Smart & Good Schools Test</i>
	To what extent did students perceive the transitional environment?	<i>High School Transition Survey (HSTS)</i>
	On an ongoing basis do stakeholders perceive that they have a voice in designing the Fisd Character Education Program?	<i>Building Developmental Assets Pre-Post-test</i>

3. RESULTS

This section is a complete compilation of program evaluation results recorded objective-by-objective. Each has a brief explanation of the method(s) and measure(s), as well as their results. In some instances, depending upon the measure, the results are compiled into one “results” sub-section, and in others the results are described based on the time frame in which the evaluation investigation was conducted.

3.1 OUTCOMES OF OBJECTIVE ONE

3.1.1 To what extent did teachers’ knowledge and awareness increase?

This evaluation component follows an *explanatory design* method, whereby a qualitative study follows a quantitative study. Quantitative data were collected first by means of a survey instrument, the results of which drove the qualitative phase to explain, elucidate, and elaborate on the quantitative findings.

Phase I – Quantitative Program Evaluation

Null Hypothesis (H_{01}) – There is no difference between classroom character-trait perceptions of experimental teachers when compared to control group teachers.

Population – Floresville ISD teachers.

Sample – Proportional stratified sample. Character education teachers, $n=24$; Non-character education teachers, randomly selected, $n=24$.

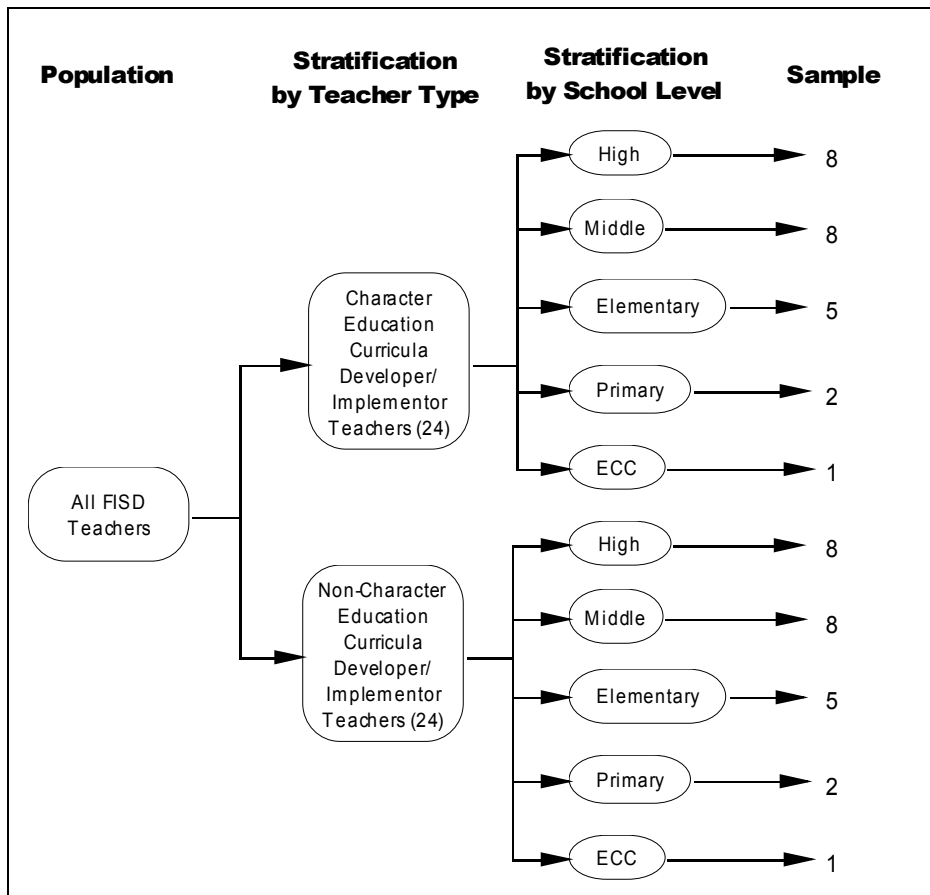


Figure 3.1.1. Stratified sample

Instrument – Modified versions of a variety of pre-existing instruments and new items to measure teachers’ perceptions of six character-trait scales of: (1) trustworthiness, (2) respect, (3) responsibility, (4) fairness, (5) caring, and (6) citizenship in their classrooms was developed as the Classroom Community Environment Survey (CCES). The CCES was administered via the Web to reduce response errors.

Unit of Analysis – Classroom teacher.

Measurement – 1 pre-treatment administration (December 2006) which is referred to as the “baseline” study, three staggered post-treatment administrations (March, August, December 2007).

Analysis – Construct validity (all responses combined), reliability (Cronbach’s alpha per administration), descriptive statistics, and inferential test (parametric) independent-samples *t*-test.

3.1.1.1 Teacher CCES 2006 Baseline

The CCES was administered in December 2006 to a control group of 24 teachers and an experimental group of 22 teachers. Cronbach’s Alpha was used to determine the reliability of each scale and overall instrument reliability. Scale reliability fell between

0.84 and 0.94, while the overall reliability was 0.97 for the entire instrument for this administration (Table 3.1.1.1a).

Table 3.1.1.1a. December 2006 CCES scale mean, scale reliability, and instrument reliability

All Teachers Responding	M	α
Trustworthiness	3.8	0.88
Respect	3.8	0.94
Responsibility	3.6	0.84
Fairness	3.6	0.90
Caring	3.7	0.92
Citizenship	3.5	0.86
N=46		0.97

When the scale mean of the experimental group and the control group are compared there was little difference between the groups (Table 3.1.1.1b).

Table 3.1.1.1b. December 2006 experimental and control group means and differences between means

	Experimental Group (n=22)	Control Group (n=24)	Difference between means
Trustworthiness	3.88	3.73	0.15
Respect	3.89	3.83	0.06
Responsibility	3.64	3.63	0.01
Fairness	3.62	3.59	0.03
Caring	3.80	3.61	0.19
Citizenship	3.51	3.48	0.03
			Mean difference 0.08

Levine's item-by-item test of heteroschedasticity was used to determine if the variance of the two groups' distributions differed significantly. At a 95% confidence level CCES Trustworthiness item 6 ($F = 5.29, p = 0.026$), Respect item 10 ($F = 4.78, p = 0.034$), Caring item 43 ($F = 7.97, p = 0.007$), and Citizenship items 48 ($F = 9.19, p = 0.004$) and 54 ($F = 5.87, p = 0.020$) were found to have significant variances and on the following Independent-Samples t -test unequal-variance estimates were used rather than the equal-variance estimates for these items. The results of the Independent-Samples t -test indicated that means of the experimental group and the control group did not differ significantly at the 95% confidence level on any items except Trustworthiness item 7 ($t = -2.09, df = 44, p = 0.042$) when equal variances were assumed and Fairness item 28 ($t = 2.05, df = 44, p = 0.046$) when equal variances were assumed. Cohen's $d = 0.3$.

3.1.1.2 Teacher CCES 2007 Second Administration

The CCES was administered a second time three months after baseline data were collected, after teacher professional development, and after the experimental group was tasked with developing "character education" lessons and teaching them. Twenty-one of the original 22 experimental teachers participated and all 24 of the randomly-selected control group participated in this second administration.

Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability of the second administration of the CCES. Scale reliability fell between 0.82 and 0.98, while the overall reliability was 0.98 for the entire instrument for this administration (Table 3.1.1.2a).

Table 3.1.1.2a. March 2007 CCES scale mean, scale reliability, and instrument reliability

All Teachers Responding	M	α
Trustworthiness	3.84	0.91
Respect	3.89	0.92
Responsibility	3.7	0.82
Fairness	3.69	0.90
Caring	3.81	0.95
Citizenship	3.67	0.92
Instrument		0.98

N = 45

When the scale mean of the experimental group and the control group were compared there was little difference between the groups (0.19) (table 3.1.1.1b), however, the difference was greater than reported in the baseline CCES administration (0.08).

Table 3.1.1.2b. March 2007 experimental and control group means and differences between means

	Experimental Group (n=21)	Control Group (n=24)	Difference
Trustworthiness	3.97	3.77	0.20
Respect	4.07	3.78	0.29
Responsibility	3.79	3.70	0.09
Fairness	3.78	3.64	0.14
Caring	3.98	3.71	0.27
Citizenship	3.78	3.58	0.20
			Mean diff. = 0.19

Levine's item-by-item test of heteroschedasticity was used to determine if the variance of the two groups' distributions differed significantly. At a 95% confidence level CCES Trustworthiness item 7 ($F = 5.76$, $p = 0.023$) and Citizenship item 53 ($F = 6.88$, $p = 0.014$) were found to have significant variances. Therefore, on the Independent-Samples t -test unequal-variance estimates were used rather than the equal-variance estimates for the above items.

The results of the Independent-Samples t -test indicated that means of the **experimental group and the control group differed significantly at the 95% confidence level on 15 items** (Table 3.1.1.2c).

Table 3.1.1.2c. Independent-samples t -test results with significant differences.

Scale	Item Number	t	p
Trustworthiness	3	2.7	0.010
	6	2.3	0.023
Respect	10	2.5	0.018
	11	2.6	0.014
	12	2.4	0.020
	13	2.9	0.006
	14	3.0	0.005

	15	3.0	0.004
Responsibility	19	-2.1	0.038
Fairness	37	2.1	0.043
Caring	39	2.9	0.006
	44	2.1	0.044
	46	2.7	0.010
	50	2.5	0.018
Citizenship	52	2.4	0.019

$p < 0.05$, $df = 29$

These results indicate that the experimental group perceived more character education characteristics on six CCES scales than did the control group, with the most changes related to teachers' perceptions of student Respect (6) and Caring (4). However, control group teachers had stronger perceptions related to student Responsibility than did the experimental teachers. Effect size remained small at Cohen's $d = 0.1$.

The overall the null hypothesis...

Null Hypothesis (H_{01}) – There is no difference between classroom character-trait perceptions of experimental teachers when compared to control group teachers.

...cannot be supported completely based on the results of the CCES baseline and second administration on certain items and on the scales of Caring and Respect. An alternative hypothesis can be suggested that there is a significant difference in character-trait perceptions of experimental teachers when compared to control group teachers on 15 Classroom Community Environment Survey (CCES) items (Table 3.1.1.2c).

3.1.1.3 2006-2007 CCES Descriptive Comparison

Table 3.1.1.3 presents an overall increase in all teachers (control and experimental) perceptions related to student character on six scales from baseline to the second administration of the CCES.

Table 3.1.1.3. CCES baseline (2006) and second administration (2007) changes in scale mean.

All Teachers Responding	2006	2007	Difference
Trustworthiness	3.81	3.84	+0.03
Respect	3.84	3.89	+0.05
Responsibility	3.33	3.4	+0.07
Fairness	3.61	3.69	+0.08
Caring	3.7	3.81	+0.11
Citizenship	3.5	3.67	+0.17
	N=46	N=45	Mean +0.9

3.1.1.4 Teacher CCES 2007 Third Administration

The CCES was administered a third time eight months after baseline data were collected and five months after the second administration of the instrument. Fourteen of the original 22 experimental teachers participated and 24 randomly-selected control group teachers from another school district participated in this third administration.

Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability of the second administration of the CCES. Scale reliability fell between 0.82 and 0.98, while the overall reliability was 0.98 for the entire instrument for this administration (Table 3.1.1.4a).

Table 3.1.1.4a. August 2007 CCES scale mean, scale reliability, and instrument reliability.

All Teachers Responding	M	α
Trustworthiness	3.83	0.94
Respect	4.05	0.93
Responsibility	3.68	0.88
Fairness	3.61	0.92
Caring	3.73	0.90
Citizenship	3.50	0.90
Instrument		0.97

N = 38

When the scale means of the experimental group and the control group were compared there was little difference between the groups (0.07) (Table 3.1.1.4b), similar to the difference in means reported in the baseline CCES administration (0.08).

Table 3.1.1.4b. August 2007 experimental and control group means and differences between means.

	Experimental Group (n=14)	Control Group (n=24)	Difference
Trustworthiness	3.88	3.82	0.06
Respect	4.00	3.90	0.10
Responsibility	3.32	3.34	-0.02
Fairness	3.67	3.56	0.11
Caring	3.81	3.56	0.25
Citizenship	3.56	3.50	0.06
			Mean diff. = 0.07

Levine's item-by-item test of heteroschedasticity was used to determine if the variance of the two groups' distributions differed significantly. At a 95% confidence level CCES Trustworthiness item 3 ($F = 5.67, p = 0.023$), Respect item 18 ($F = 10.94, p = 0.002$), and Citizenship items 42 ($F = 4.82, p = 0.035$) and 45 ($F = 7.95, p = 0.008$) were found to have significant variances. Therefore, on the Independent-Samples t -test unequal-variance estimates were used rather than the equal-variance estimates for the above items.

The results of the Independent-Samples t -test indicated that means of **the experimental group and the control group differed significantly at the 95% confidence level on 4 items** (Table 3.1.1.4c).

Table 3.1.1.4c. Independent-samples t -test results with significant differences.

Scale	Item Number	<i>t</i>	<i>p</i>
Trustworthiness	6	2.1	0.046
Respect	22	2.8	0.009
Caring	42	2.1	0.046
Citizenship	53	2.0	0.049

$p < 0.05$, $df = 36$

These results of the Independent-Samples *t*-test indicated that means of the experimental group and the control group did not differ significantly at the 95% confidence level on any items except Trustworthiness item 7, Respect item 22, Caring item 42, and Citizenship item 53. Effect size remained small at Cohen's $d = 0.21$.

3.1.1.5 CCES Descriptive Comparison of Repeated Measures

Table 3.1.1.5a presents changes in all teachers' (control and experimental) perceptions related to student character on six scales in a repeated measures experimental method (i.e. baseline administration, second administration, and third administration of the CCES).

Table 3.1.1.5a. CCES baseline, second, and third administration changes in scale mean.

All Teachers Responding	Baseline	2 nd Administration	3 rd Administration
Trustworthiness	3.81	3.84	3.83
Respect	3.84	3.89	4.05
Responsibility	3.33	3.40	3.68
Fairness	3.61	3.69	3.61
Caring	3.70	3.81	3.73
Citizenship	3.50	3.67	3.50
	<i>N</i> =46	<i>N</i> =45	<i>N</i> =38

Table 3.1.1.5b presents changes in all teachers' perceptions when results are divided in to experimental and control groups. Likewise, Figure 3.1.1.5 presents the same data in a graphic format. What is noticeable is that there is little practical significance in teachers' perceptions on six character education scales. Generally, the experimental group appears to have stronger perceptions of their students compared to control group teachers, however, the differences between groups are negligible from a practical significance viewpoint.

Table 3.1.1.5b. CCES baseline, second, and third administration changes in scale mean.

	Baseline		2 nd Administration		3 rd Administration	
	Experimental Group (n=22)	Control Group (n=24)	Experimental Group (n=21)	Control Group (n=24)	Experimental Group (n=14)	Control Group (n=24)
Trustworthiness	3.88	3.73	3.97	3.77	3.88	3.82
Respect	3.89	3.83	4.07	3.78	4.00	3.90
Responsibility	3.64	3.63	3.79	3.70	3.32	3.34
Fairness	3.62	3.59	3.78	3.64	3.67	3.56
Caring	3.80	3.61	3.98	3.71	3.81	3.56
Citizenship	3.51	3.48	3.78	3.58	3.56	3.50

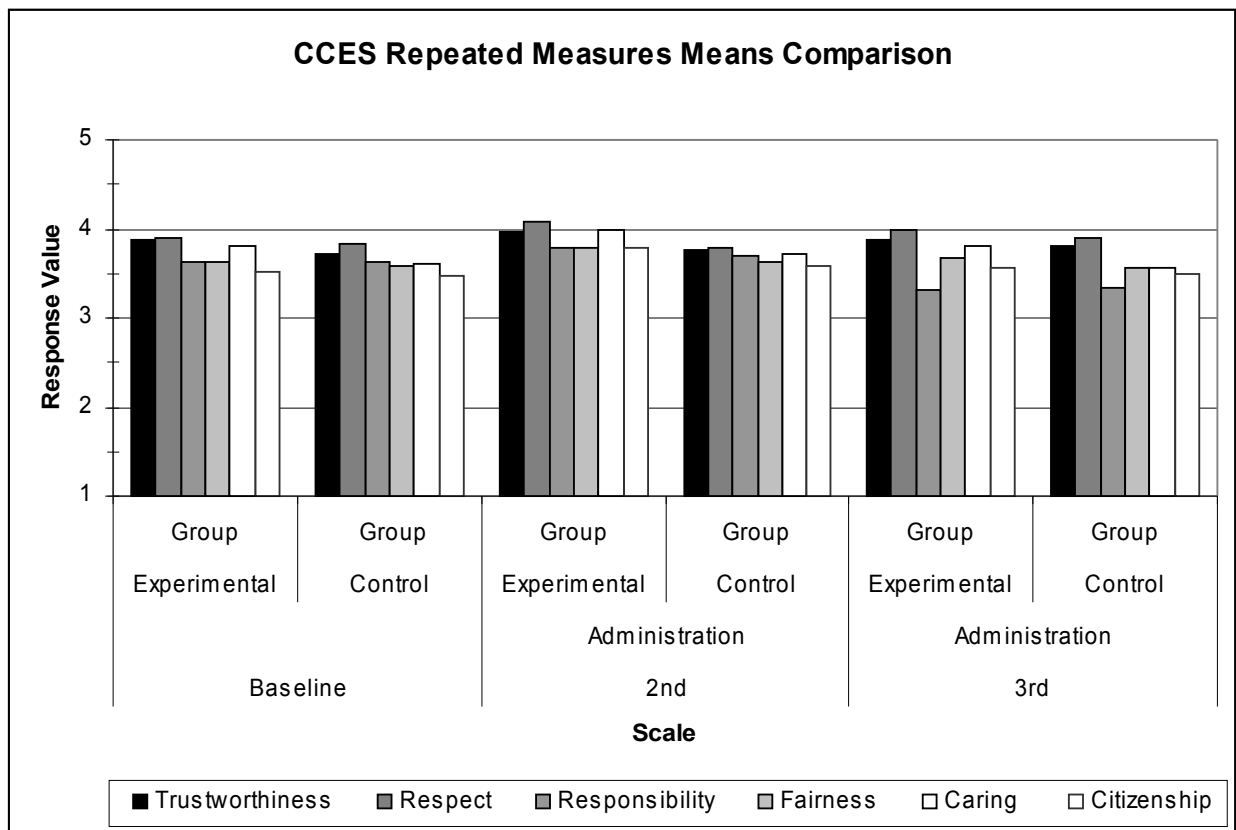


Figure 3.1.1.5. CCES baseline, second, and third administration changes in scale mean

3.1.1.6 CCES Inferential Analyses of Repeated Measures

Paired-samples *t*-tests were conducted on data from the experimental group baseline, first, and second administrations of the CCES for teachers (Table 3.1.1.6a).

Table 3.1.1.6a. Paired-samples *t*-tests and correlations of experimental group perceptions over time

Pair/Scale	<i>N</i>	<i>r</i>	<i>p</i>	<i>t</i>	<i>df</i>	<i>p</i>
Baseline to Second Administration						
Trustworthiness	23	.846	.000*	-1.006	22	.325
Respect	23	.764	.000*	-1.914	22	.069
Responsibility	23	.660	.001*	-1.141	22	.266
Fairness	23	.825	.000*	-1.777	22	.089
Caring	23	.713	.000*	-1.444	22	.163
Citizenship	23	.758	.000*	-2.916	22	.008*
Baseline to Third Administration						
Trustworthiness	14	.461	.097	.513	13	.617
Respect	14	.176	.547	-.063	13	.951
Responsibility	14	.016	.956	.590	13	.565
Fairness	14	.175	.550	.819	13	.427
Caring	14	.305	.288	.043	13	.967
Citizenship	14	.242	.404	.048	13	.963

**p*<0.05

Results from the baseline to the second administration demonstrate substantial and significant correlations on all scales. In other words, the treatment effect was substantial. Likewise, the mean perception on the baseline administration was significantly greater than the mean perceptions on the second administration. In contrast, when considering data from the baseline to the third administration of the CCES, there is little correlation on all scales. The mean perceptions on the third administration were not significantly greater than the mean perceptions during the baseline administration. Put plainly, nine months after the baseline administration and eight months after character education professional development, there is no evidence in the data that teachers' perceptions of student character exhibited a difference than before treatment.

When control group data are analyzed the results are similar to that of the experimental group. Table 3.1.1.6b presents paired-samples *t*-test results from the baseline to the second administration of the CCES with control group data only. It also presents results from an independent-samples *t*-test from baseline to the third administration of the CCES. The control group changed (see 3.1.1.7 Teacher CCES 2007 Fourth Administration for details) from the second to the third administration of the instrument, thus the change from a paired-samples test to an independent samples test.

Table 3.1.1.6b. *T*-tests and correlations of the control group perceptions over time. **p*<0.05

Pair/Scale	<i>N</i>	<i>r</i>	<i>p</i>	<i>t</i>	<i>df</i>	<i>p</i>
Baseline to Second Administration Paired-Samples <i>t</i> -test						
Trustworthiness	24	.923	.000*	-.345	23	.734
Respect	24	.592	.002*	.295	23	.771
Responsibility	24	.823	.000*	.000	23	1.00
Fairness	24	.768	.000*	-.904	23	.375
Caring	24	.634	.001*	-.648	23	.524
Citizenship	24	.727	.000*	-.068	23	.947
Baseline to Third Administration Independent-Samples <i>t</i> -test						
Trustworthiness				1.433	47	.175
Respect				-1.139	47	.275
Responsibility				-1.139	47	.275
Fairness				.958	47	.356
Caring				.380	47	.710
Citizenship				-.145	47	.887

From the baseline to the second administration of the CCES there are substantial and significant correlations on all scales. Despite no treatment, the control group teachers' mean perception of their students' character was high during both the baseline administration and the second administration. This result is likely explained by diffusion of treatment corrupting experimental validity. On the other hand, when a new out-of-school district control group was used, from the baseline administration to the third administration, the scale means did not differ significantly.

When an analysis of variance was completed on data from the baseline and second administrations of the CCES there were no statistically significant differences between the experimental group and the control group on six scales in most instances. There were significant differences in teachers' perceptions at the 95% confidence level on the scales of Caring and Citizenship on the second administration of the CCES (Table 3.1.1.6c).

Table 3.1.1.6c.. One-way ANOVA of two administrations of the CCES

Scale	Baseline (N=46)		Second Administration (N=45)	
	F	p	F	p
Trustworthiness	2.96	.092	4.04	.050
Respect	0.96	.330	5.75	.021
Responsibility	1.56	.218	4.30	.044
Fairness	1.05	.309	2.57	.115
Caring	3.67	.062	5.86	.020*
Citizenship	1.19	.279	7.52	.009*

*p<0.05

3.1.1.7 Teacher CCES 2007 Fourth Administration

The CCES was administered a fourth time 12 months after baseline data were collected, 10 months after the second administration, and four months after the third administration. Results from the 2006-2007 school year can only be casually compared to the results from 2007-2008 school year because the control group changed in the second year (i.e. 2007-2008). Originally, the control group consisted of Floresville Independent School District (FISD) teachers. However, given FISD's small size, the control group was "washing out," or experiencing *diffusion of treatment* internal validity problems. In other words, it became impossible for the FISD control group to NOT be exposed to character education initiatives. The program evaluator took the advice of other PCEP grant evaluators at the January 25-26, 2007, PCEP Grant Evaluator meeting in Orange, CA and suggested the control group be selected from another school district to remediate the situation. FISD grant administrators were able to obtain the cooperation of teachers in Stockdale ISD and Poth ISD, both neighboring, small, rural school districts with similar teacher/student populations and no character education initiatives. Therefore, the control group for the following 2007-2008 results differs from CCES baseline and second administrations. The selection method for the new out-of-district control group was identical to that of the first in-district control group. That is, the new control group was selected randomly from a stratified teacher population.

Twenty five experimental teachers and 32 randomly-selected control group teachers participated in this fourth CCES administration. Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability of the second administration of the CCES. Scale reliability fell between 0.59 and 0.94, while the overall reliability was 0.97 for the entire instrument for this administration (Table 3.1.1.7a).

Table 3.1.1.7a. December 2007 CCES scale mean, scale reliability, and instrument reliability.

All Teachers Responding	M	α
Trustworthiness	3.82	0.93
Respect	3.79	0.94
Responsibility	3.32	0.59
Fairness	3.55	0.90
Caring	3.68	0.91
Citizenship	3.45	0.88
Instrument		0.97

N = 57

When the scale means of the experimental group and the control group were compared there was little difference between the groups (0.3) (Table 3.1.1.7b).

Table 3.1.1.7b. August 2007 experimental and control group means and differences between means.

	Experimental Group (<i>n</i> =25)	Control Group (<i>n</i> =32)	Difference
Trustworthiness	3.8	3.8	0.0
Respect	3.9	3.4	0.5
Responsibility	3.4	3.2	0.2
Fairness	3.7	3.4	0.3
Caring	3.8	3.6	0.2
Citizenship	3.6	3.3	0.3
			Mean diff. = 0.3

Levine's item-by-item test of heteroschedasticity was used to determine if the variance of the two groups' distributions differed significantly. At a 95% confidence level CCES Citizenship items 50 ($F = 4.13$, $p = 0.047$), 52 ($F = 15.12$, $p = 0.000$), and 54 ($F = 4.62$, $p = 0.036$) were found to have significant variances. Therefore, on the Independent-Samples *t*-test unequal-variance estimates were used rather than the equal-variance estimates for those items.

The results of the Independent-Samples *t*-test indicated that means of **the experimental group and the control group differed significantly at the 95% confidence level on 12 items** (table 3.1.1.4c), the majority of which (5) fell on the scale of Citizenship. Cohen's $d = 0.32$.

Table 3.1.1.7c. Independent-samples *t*-test results with significant differences.

Scale	Item Number	<i>t</i>	<i>p</i>
Respect	18	2.102	.040
	21	2.215	.031
Responsibility	22	2.046	.046
	25	2.537	.014
	28	3.215	.002
Fairness	36	2.143	.037
	46	2.561	.013
Caring	50	2.267	.029
	52	2.675	.011
	53	2.618	.011
	54	2.024	.049
Citizenship	55	2.941	.005

p < 0.05, *df* = 55

When an analysis of variance was completed on data from the third (August 2007) and fourth (December 2007) administrations of the CCES there were no statistically significant differences between the experimental group and the control group on six scales (Table 3.1.1.7d).

Table 3.1.1.7d.. One-way ANOVA of two administrations of the CCES

Scale	Baseline (N=35)		Second Administration (N=45)	
	<i>F</i>	<i>p</i>	<i>F</i>	<i>p</i>
Trustworthiness	.087	.770	0.029	.866
Respect	.061	.806	3.389	.075
Responsibility	.053	.819	2.649	.113
Fairness	.031	.862	0.741	.395
Caring	.002	.966	2.369	.133
Citizenship	.001	.974	1.658	.207

**p*<0.05

3.1.1.8 Time Period Fall 2007 Teacher CCES

The Classroom Community Environment Survey (CCES) was administered in August 2007 and December 2007. In order to examine the relationships that may exist among the subscales correlation analyses were conducted. The results indicate poor associations among the six scales of Trustworthiness, Respect, Responsibility, Fairness, Caring, and Citizenship (August 2007 *r* < .45, December 2007 *r* < .51).

Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability of the August administration of the CCES. Scale reliability fell between 0.82 and 0.98, while the overall reliability was 0.98 for the entire instrument for this administration (Table 3.1.1.8a).

Table 3.1.1.8a. August 2007 CCES scale mean, scale reliability, and instrument reliability

All Teachers Responding	M	α
Trustworthiness	3.83	0.94
Respect	4.05	0.93
Responsibility	3.68	0.88
Fairness	3.61	0.92
Caring	3.73	0.90
Citizenship	3.50	0.90
Instrument		0.97

N = 38, Response scale = 1-Almost Never 2-Rarely 3- Sometimes 4-Often 5-Almost Always

Scale reliability fell between 0.59 and 0.94, while the overall reliability was 0.97 for the entire instrument for the December administration (Table 3.1.1.8b).

Table 3.1.1.8b. December 2007 CCES scale mean, scale reliability, and instrument reliability

All Teachers Responding	M	α
Trustworthiness	3.82	0.93
Respect	3.79	0.94
Responsibility	3.32	0.59
Fairness	3.55	0.90
Caring	3.68	0.91
Citizenship	3.45	0.88
Instrument		0.97

N = 57, Response scale = 1-Almost Never 2-Rarely 3- Sometimes 4-Often 5-Almost Always

A MANOVA with the August and December 2007 administrations as independent variables (Time) and experimental group Trustworthiness, Respect, Responsibility, Fairness, Caring, and Citizenship as the dependent variables demonstrated no statistically significant effect [$F(11, 6) = 1.0, p = 1.00, \lambda = 1.0, \text{partial } \eta^2 = .00, \text{power} = .05$]. Examination of experimental group results vs. control group results indicated no statistically significant main effect [$F(2, 12) = 2.35, p = .178, \lambda = .151, \text{partial } \eta^2 = .85, \text{power} = .38$].

3.1.1.9 CCES Construct Validity

Validity was considered in terms of content or 'face' validity and also in terms of construct validity during the development of the Classroom Community Environment Survey (CCES). Construct validity was investigated as described below using principal component factor analysis with varimax rotation and Kaiser normalization. The aim of factor analysis is to ascertain the fundamental structure of a comparatively large set of variables (Garson, 2001). This method of analysis is consistent with the intuitive-rational method of learning environment instrument development (Fraser, 1986) and has been used (Dorman, 2003; Walker, 2003) to determine if items load on a priori scales. In essence, factor analysis provides information about whether items within a given scale are measuring that scale and no other scale. Only those items with a factor loading of at least 0.40 with their own scale were kept in the refined CCES.

Table 3.1.1.9 presents the rotated component matrix based on individual items. No 'faulty' items of the original 55 items were identified. Six scales were originally developed for the CCES and after factor analysis the same six scales remained: *Trustworthiness, Respect, Responsibility, Fairness, Caring, and Citizenship.*

Table 3.1.1.9. Rotated component matrix factor loadings for the 55-item CCES

	Trust	Respect	Responsibility	Fairness	Caring	Citizenship
Trustworthiness 1	.739					
Trustworthiness 2	.624					
Trustworthiness 3	.604					
Trustworthiness 4	.642					
Trustworthiness 5	.602					
Trustworthiness 6	.506					
Trustworthiness 7	.457					
Trustworthiness 8	.661					
Trustworthiness 9	.637					
Respect 10		.738				
Respect 11		.740				
Respect 12		.717				
Respect 13		.759				
Respect 14		.785				
Respect 15		.638				
Respect 16		.708				
Respect 17		.597				
Respect 18		.697				
Responsibility 19			.447			
Responsibility 20			.712			
Responsibility 21			.521			
Responsibility 22			.710			
Responsibility 23			.411			
Responsibility 24			.531			
Responsibility 25			.462			
Responsibility 26			.661			
Responsibility 27			.451			
Fairness 28				.815		
Fairness 29				.639		
Fairness 30				.653		
Fairness 31				.575		
Fairness 32				.658		
Fairness 33				.591		
Fairness 34				.672		
Fairness 35				.579		
Fairness 36				.424		
Fairness 37				.597		
Caring 38					.617	
Caring 39					.587	
Caring 40					.668	
Caring 41					.394	
Caring 42					.640	
Caring 43					.744	
Caring 44					.511	
Caring 45					.570	

Caring 46	.552	
Citizenship 47		.400
Citizenship 48		.545
Citizenship 49		.552
Citizenship 50		.576
Citizenship 51		.490
Citizenship 52		.510
Citizenship 53		.520
Citizenship 54		.440
Citizenship 55		.659

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Factor loadings smaller than 0.40 have been omitted. N=139

When considering discriminant validity through factor analysis, each scale should measure a unique dimension not measured by another scale. Discriminant validity can be improved by removing any item whose factor loading falls below the 0.40 threshold within its a priori assigned scale or was above 0.40 with any of the other five scales. In this case no items needed to be removed to improve discriminant validity.

Further research replicating this study should be conducted in the future with larger and different samples to demonstrate the replication of similar results, because any given sample will never perfectly reflect its population. While deviations might be reduced with increases in sample size, deviations will still occur (Kim & Mueller, 1978) even with larger samples. Therefore, it will be important to replicate the factor analysis procedures each time the CCES is used in order to demonstrate that the factor structure presented here is not peculiar to this sample.

3.1.2 Phase II – Qualitative Program Evaluation Results

Study period December 2006 to May 2007

On May 22, 2007, four character education “experimental group” teachers were randomly selected and four control group teachers were randomly selected to participate in two separate focus group sessions. Each group consisted of one Primary, Elementary, Middle, and High school teacher. Each session was approximately one hour long. Participants were shown a bar graph (Figure 3.1.2) of the Classroom Community Environment Survey (CCES) combined results and asked as a group to speculate on particular scales. The scales of focus were Respect, Caring, and Citizenship since the response differences between groups were most evident on these scales. Each of these scales and focus group comments are presented below.

Respect

The scale of Respect had rising means from the base line to the second administration of the CCES from the experimental group, yet had reducing means from the control group (Figure 3.1.2). Teachers commented as follows regarding why they thought this was the case.

Experimental group comments

“Respect and responsibility are a continual focus in our school and are a part of our everyday curriculum.” [Primary teacher]

“In the elementary school this [Respect] is a weekly pillar and common vocabulary.”

“We identify a concept and make it explicit in the middle school.”

“We’re slow off the block in High School” [speculating why the mean for this scale is reducing at the High School level for the control group]

Control group comments

“Students don’t know any better at the primary level.”

“Elementary students are still ‘testing the waters.’”

“We don’t have time or the familiarity to specifically teach respect.” [Middle school teacher]

“Students have an ‘I don’t care’ attitude at the high school level.”

Analysis – Experimental group teachers seem to simply assume character education is just a part of their day-to-day teaching, however, the control group teachers tended to place blame for their perceptions lowering on the students or a lack of time to add character education to the curriculum (viewing character education as a add on).

Caring

The scale of Caring in the bar graph shown to teachers (Figure 3.1.2) presents a rising mean among experimental group teachers’ perceptions and a slightly rising, yet lower mean among control group teachers’ perceptions. The control group teachers tended to speculate more on the reason behind this as noted below.

Experimental group comments

“The trained teachers [meaning the experimental group] see [perceive] this more because it’s labeled.” [Elementary teacher]

“High school teachers may think it’s up to students to care” [speculating on why the control group has lower means.]

Control group comments

“Primary students have few opportunities to show caring.”

[Primary students] “may get frustrated, especially toward the end of the school year when this [survey] was done.”

“It’s a result of their home environment.” [Elementary teacher]

“Sixth graders don’t really care. Some do, but it stems from home.”

“In our high school these students have grown up with each other and familiarity may lead to it being easier for them not to care.”

Analysis – Control group teachers tended to place the direction of student caring on external factors such a home environment and the students themselves. While the experimental group had fewer comments, the Elementary teachers’ speculation that because they are aware of Caring it naturally leads to noticing it more.

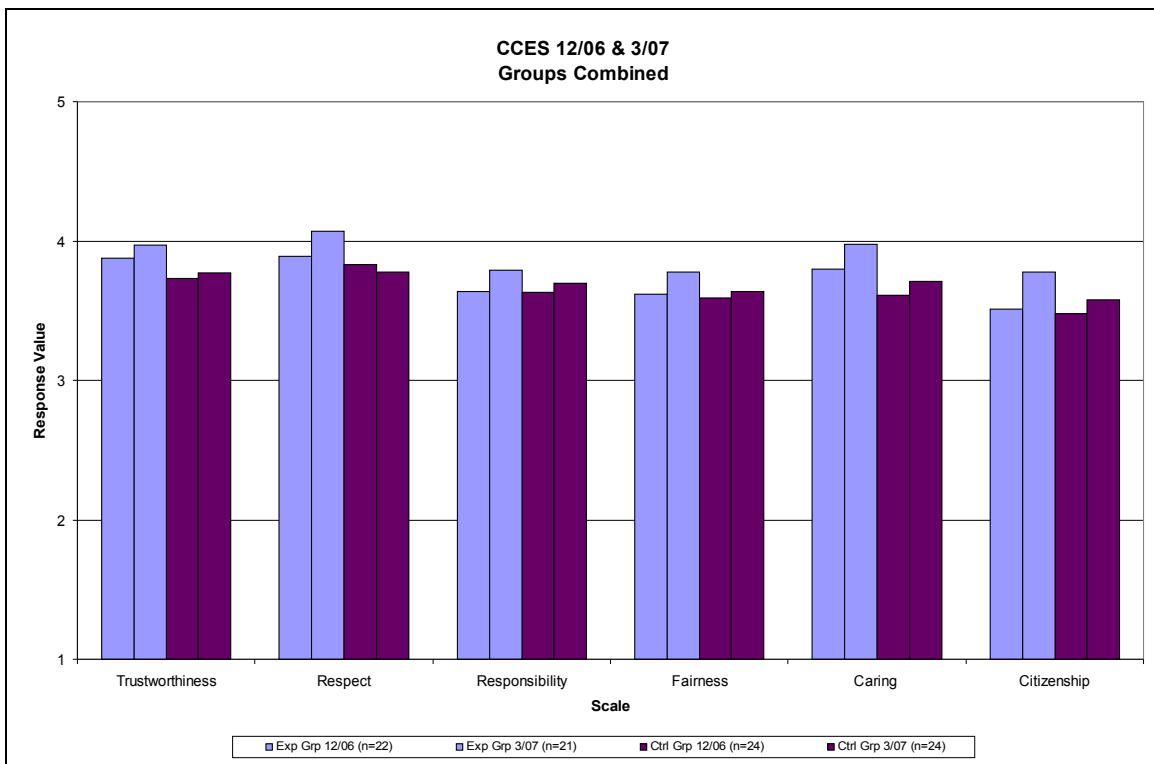


Figure 3.1.2. Classroom Community Environment Survey (CCES) results

Citizenship

Like the scale of Caring, the scale of Citizenship indicated a rise in experimental group teachers’ perceptions of it in their students. The control group teachers noticed Citizenship more in May, but still less so than the experimental group teachers. When asked about this scale the experimental group had little to say about why their perceptions rose. Their comments were directed more at why the control group had lower perception means.

Experimental group comments

“This is a difficult concept for elementary students.”

“Even high school students have difficulty understanding this [citizenship].

“This is based on the individual teacher. Some middle school teachers don’t make their students stand for the Pledge [of Allegiance], others do.”

“Same thing at the elementary school [referring to middle school teacher comment]. Some teachers use this time [Pledge of Allegiance] to teach about citizenship, others catch up on work at their desk.”

Control group comments

“It [citizenship] must be backed up by the community and parents.” [High school teacher]

“Yes, there are mixed messages. The dress code for example. Some teachers enforce it, others don’t care.” [Middle school teacher]

Analysis – Both groups of teachers seemed to acknowledge that the concept of citizenship is difficult to understand. Even with the Citizenship survey items in hand, the teachers focused on the Pledge of Allegiance, a recent school discipline situation involving a High School athlete, and the dress code rather than what the survey indicated related to resolving conflicts—taking an active role in governance, and working together. Perhaps teachers themselves have difficulty understanding the concept of citizenship when placed in this context of character education.

Study period August 2007 to April 2008

A Phase II qualitative follow-up was conducted on April 16, 2008. Experimental group primary/elementary teachers ($n=10$), middle school teachers ($n=7$), and high school teachers ($n=4$) participated in three separate focus group sessions. Participants were shown bar graphs of a variety of Classroom Community Environment Survey (CCES) results and asked as a group to speculate on particular scale anomalies as well as 2006 – 2007 data trends. They were also asked to identify process issues of (1) program barriers, (2) program things that work well, and (3) possible program changes that could improve character education as at the school level.

All teachers – When asked about the trend of falling perceptions of student character on six scales from 2006 to present (Figure 3.1.2b), teachers attributed them to their rising expectations of students (thus, teachers are more aware of character education issues in their classrooms and don’t rank their students as strong as in the early part of the

grant program) and “survey fatigue” (i.e. constantly doing surveys for this program and others).

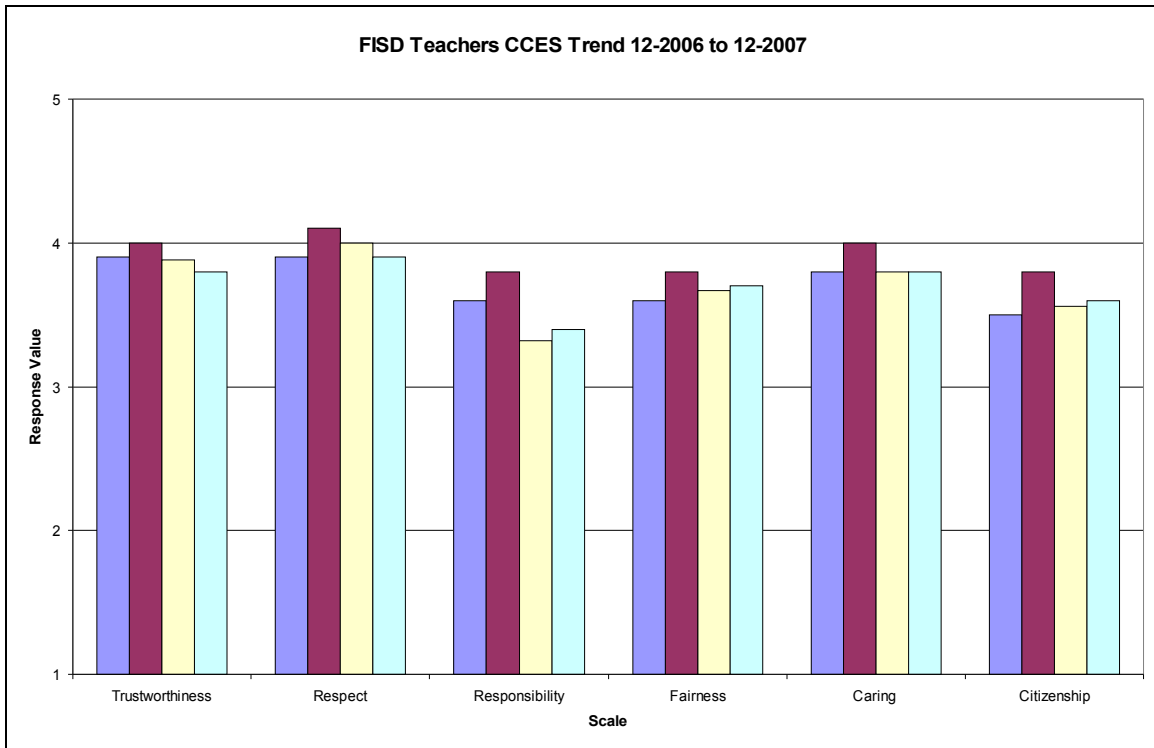


Figure 3.1.2b. All FISD teacher CCES results

Primary/Elementary teachers – When asked about a drop in perceptions related to the scale of Trustworthiness from 8-2007 to 12-2007, teachers stated that as students become more comfortable in their classes they are less inhibited by teachers and fellow students, thus they tend to push truth and honesty-oriented characteristics further.

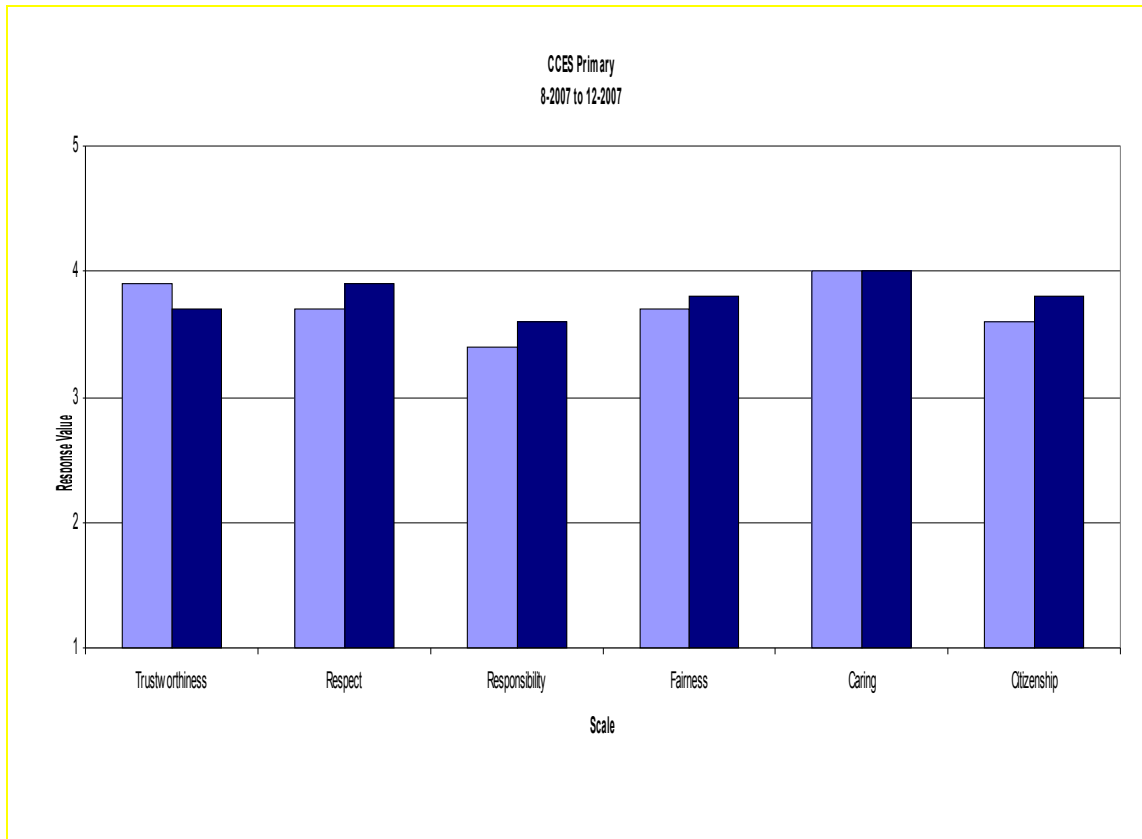


Figure 3.1.2c. Primary/elementary school teacher CCES results

Middle school teachers – When asked about a drop from 8-2007 to 12-2007 on the scale of Respect teachers attributed it to the nearing holiday season. They stated that a similar phenomenon happens just before spring break and the end of the school year. Further, when shown a bar chart of middle school student CCES results compared to middle school teacher CCES results where there is a noticeable difference between teacher/student perceptions on the scales of Trustworthiness and Fairness (teachers perceiving stronger character than the students), teachers noted that students know what is “really going on” related to trust in the school. Likewise, teachers attributed their stronger perceptions of student fairness as related to their limited classroom perspectives as teachers. They think students may know more about “hallway behavior” being less fair than the teachers’ view in the classroom.

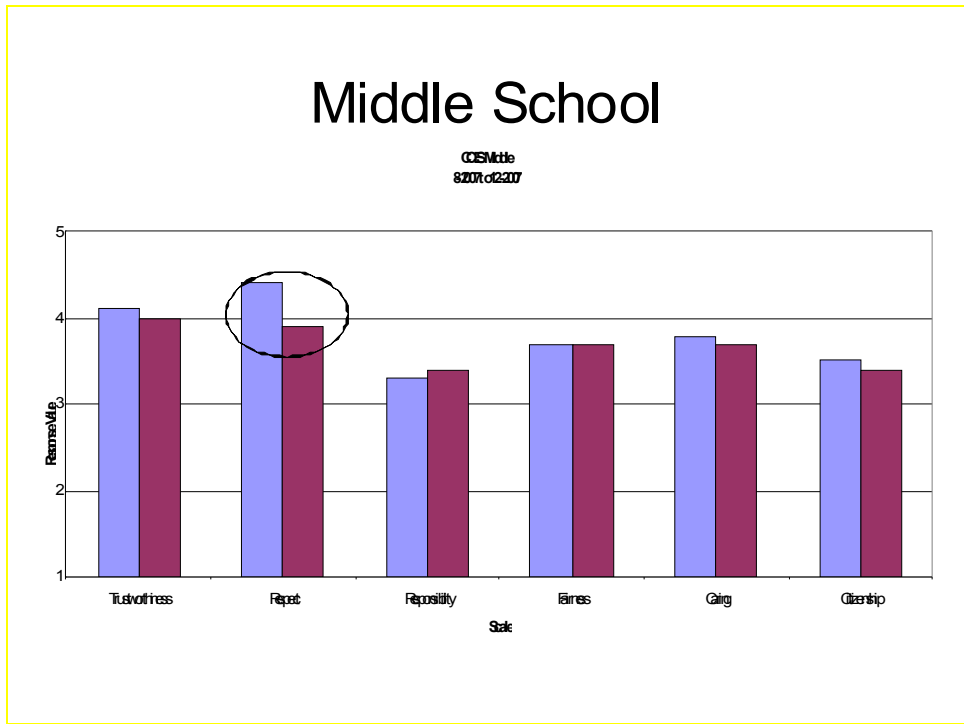


Figure 3.1.2d. Middle school teacher CCES results

High school teachers – When asked about a sharp rise in their perspectives related to the scale of Responsibility from 8-2007 to 12-2007, teachers attributed it to students becoming familiar with teacher expectations over the course of the school year. Also, CCES teacher results and CCES student results were presented. On the scales of Respect and Citizenship teachers had stronger perceptions than their students. The high school teachers, like the middle school teachers, attributed this difference to the fact that they don't see what goes on in the hallways and before/after school related to students respecting each other and teachers, and students “doing the right thing” as found in the Citizenship scale items.

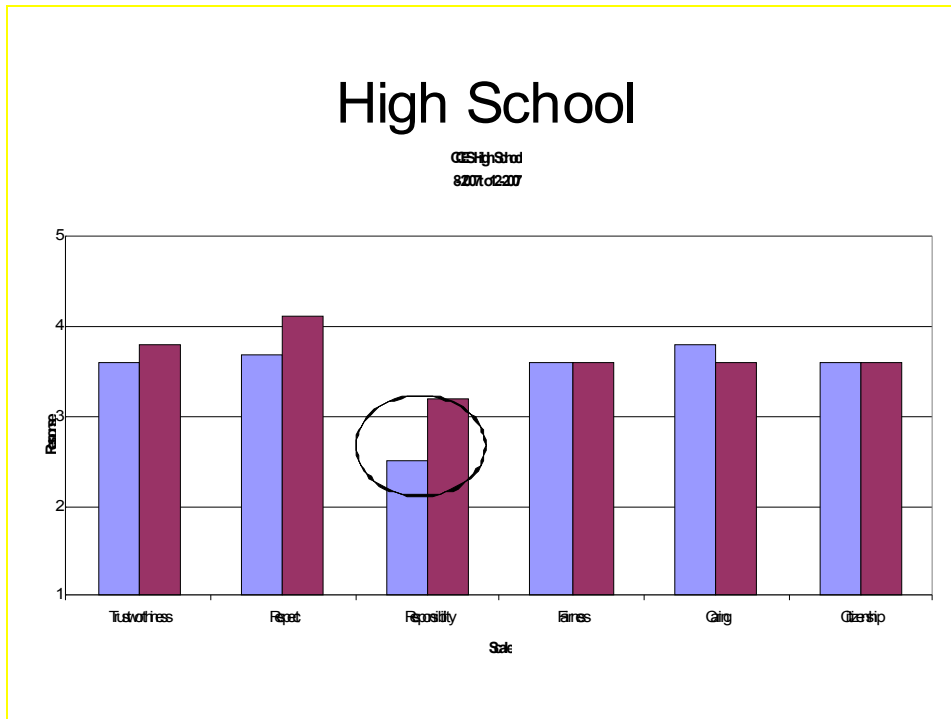


Figure 3.1.2e. High school teacher CCES results

3.1.3 To what extent did students internalize the six character trait?

This evaluation component follows an *explanatory design* method whereby a qualitative study follows a quantitative study. Quantitative data were collected first by means of a survey instrument, the results of which drove the qualitative phase to explain, elucidate, and elaborate on the quantitative findings.

3.1.3.1 Phase I – Quantitative Program Evaluation

Null Hypothesis (H_{01}) – There is no difference between classroom character-trait perceptions of experimental students when compared to control group students.

Population – Floresville ISD students (experimental group) and Stockdale ISD students (control group).

Sample – The sample was randomly selected from a list of teacher names in each school. The students of these teachers were then administered the instrument.

Instrument – Modified scales from a variety of pre-existing instruments and new items to measure students' perceptions of six character-trait scales of: (1) trustworthiness, (2) respect, (3) responsibility, (4) fairness, (5) caring, and (6) citizenship in their classrooms was developed as the Classroom Community Environment Survey (CCES). The CCES was administered via the Web to reduce response errors. The CCES was customized for three levels of students: (1) elementary, with 18 items, three items per scale, (2) middle school, with 54 items, and a (3) high school version with 54 items.

Unit of Analysis – School.

Measurement – Baseline administration was conducted in November 2007, followed by post-treatment administrations in February and March 2008 for each school level (i.e. elementary, middle, and high school).

Analysis – Content validity, reliability (Cronbach’s Alpha per administration), descriptive statistics, and inferential test (parametric) independent-samples *t*-test and MANOVA were conducted for each level of this measure.

3.1.3.1.1 Elementary-Level CCES 2007 - 2008

The CCES was administered in November 2007 to a control group consisting of 17 elementary school students and an experimental group of 115 students. Cronbach’s Alpha was used to determine the reliability of each scale and overall instrument reliability. Scale reliability fell between 0.23 and 0.57, while the overall reliability was 0.75 for the entire instrument for this administration (Table 3.1.3.1.1a).

Table 3.1.3.1.1a. November 2007 CCES scale mean, scale reliability, and instrument reliability

All Students Responding	M	α
Trustworthiness	2.7	0.23
Respect	2.6	0.37
Responsibility	2.8	0.54
Fairness	3.0	0.48
Caring	3.1	0.57
Citizenship	2.8	0.37
N=132		0.75

Response scale = 1-4.

In order to examine the relationships that may exist among the subscales correlation analyses were conducted for three administration periods (baseline, second, and third). The results indicate a moderate statistically significant degrees of associations exist among the scales of Trustworthiness, Fairness, Caring, and Citizenship, and very weak associations between the scales of Respect and Responsibility. These results suggest that Trustworthiness, Caring, Citizenship, and Respect should be grouped as a new composite variable, while Responsibility and Fairness remain independent of one another and the new composite variable called Character Values (Tables 3.1.3.1.1b, c & d).

Table 3.1.3.1.1b. November 2007 correlation coefficients for character variables at pretest

	Trustworthiness	Respect	Responsibility	Fairness	Caring	Citizenship
Trustworthiness	1	.343	-.179	.622	.611	.773
Respect		1	-.778	.026	-.299	.126
Responsibility			1	-.202	-.050	.100
Fairness				1	.652	.778
Caring					1	.386
Citizenship						1
Mean	2.68	2.37	2.9	2.9	3.12	2.73
SD	.31	.68	.44	.23	.18	.27

$p < .01$ (2-tailed)

Table 3.1.3.1.1c. February 2008 correlation coefficients for character variables at second administration

	Trustworthiness	Respect	Responsibility	Fairness	Caring	Citizenship
Trustworthiness	1	.929	-.115	-.500	1.00(**)	.945
Respect		1	-.475	-.143	.929	.756
Responsibility			1	-.803	-.115	.217
Fairness				1	-.500	-.756
Caring					1	.945
Citizenship						1
Mean	2.63	2.47	2.67	2.73	3.00	2.73
SD	.23	.40	.50	.06	.12	.15

** Correlation is significant at the 0.01 level (2-tailed).

Table 3.1.3.1.1d March 2008 correlation coefficients for character variables at third administration

	Trustworthiness	Respect	Responsibility	Fairness	Caring	Citizenship
Trustworthiness	1	.994	-.500	-.189	.945	.945
Respect		1	-.590	-.082	.904	.904
Responsibility			1	-.756	-.189	-.189
Fairness				1	-.500	-.500
Caring					1	1.000(**)
Citizenship						1
Mean						
SD						

** Correlation is significant at the 0.01 level (2-tailed).

A MANOVA with administrations one ($n = 115$), two ($n = 92$), and three ($n = 91$) as the repeated variable and Responsibility, Fairness, and Character Values as the dependent variables demonstrated no significant effect for time ($p < 0.01$).

No control group data were available for administrations two and three, thus no meaningful repeated measures analyses between groups could be conducted over time. However, control and experimental data from administration one (baseline) was collected and the results of an Independent-Samples t -test indicated that means of the experimental group and the control group did not differ significantly at the 95% confidence level on any items except Character Values item 2 and Responsibility item 8, Fairness item 11, Character Values items 13 and 18. Cohen's $d = 0.1$.

Table 3.1.3.1.c. Independent-samples *t*-test items with significant differences

Scale/Item	Item	<i>t</i>	<i>p</i>
Character Values Item 2	Kids in my class keep their promises.	2.59	0.01
Responsibility Item 8	Kids in my class stay out of trouble.	-2.58	0.01
Fairness Item 11	Kids share with each other in my class.	2.19	0.03
Character Values Item 13	We care about each other in my class.	2.25	0.03
Character Values Item 18	We listen to each other's ideas in my class.	2.51	0.01

df = 131

3.1.3.1.2 Middle School-Level CCES 2007 - 2008

The CCES was administered in November 2007, February 2008, and March 2008 to middle and high school students. In order to examine the relationships that may exist among the subscales correlation analyses were conducted. The results indicate poor associations among the six scales of Trustworthiness, Respect, Responsibility, Fairness, Caring, and Citizenship (table 3.1.3.2a) from the combined second and third administration results.

Table 3.1.3.1.2a. February and March 2008 correlation coefficients for character variables

	Trustworthiness	Respect	Responsibility	Fairness	Caring	Citizenship
Trustworthiness	1	.067	.094	.196	-.123	-.266
Respect		1	.358	-.257	.268	.233
Responsibility			1	.214	.203	.500
Fairness				1	.112	.151
Caring					1	.371
Citizenship						1
Mean	3.3	3.3	3.2	3.2	3.3	3.0
SD	.18	.21	.22	.07	.16	.07

** Correlation is significant at the 0.01 level (2-tailed), *N* = 594

Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability administered at the middle school level. Scale reliability fell between 0.80 and 0.97, while the overall reliability was 0.97 for the entire instrument for this administration (table 3.1.3.1.2b).

Table 3.1.3.1.2b. November 2007 CCES Middle School scale mean, scale reliability, and instrument reliability

All Students Responding	<i>M</i>	α
Trustworthiness	3.5	0.86
Respect	3.6	0.87
Responsibility	3.4	0.80
Fairness	3.3	0.91
Caring	3.5	0.84
Citizenship	3.1	0.84
<i>N</i> =97		0.97

Response scale = 1-Almost Never 2-Rarely 3- Sometimes 4-Often 5-Almost Always

A MANOVA with administrations one (November 2007), two (February 2008), and three (March 2008) as independent variables (Time) and experimental group Trustworthiness, Respect, Responsibility, Fairness, Caring, and Citizenship as the dependent variables demonstrated no statistically significant effect [$F(12, 38) = 1.08, p = .403, \lambda = .556$, partial $\eta^2 = .254$, power = .51]. However, a significant univariate interaction between scale of Trustworthiness and time was found [$F(2, 24) = 3.79, p = 0.25$, partial $\eta^2 = .24$, power = .634] where November's $M = 3.5, SD = .17$; February's $M = 3.1, SD = .17$; and March's $M = 3.3, SD = .16$, indicating a "leveling off" after the initial CCES administration.

Examination of experimental group results vs. control group results indicated no statistically significant main effect [$F(2, 16) = 3.287, p = 0.099, \lambda = .113$, partial $\eta^2 = .887$, power = .516].

3.1.3.1.3 High School-Level CCES 2007 - 2008

The CCES was administered in November 2007, February 2008, and March 2008 to the experimental and control groups. Reliability was verified using the November 2007 data set of the control group consisting of 54 high school students and the experimental group of 155 students. Cronbach's Alpha was used to determine the reliability of each scale and overall instrument reliability. Scale reliability fell between 0.84 and 0.91, while the overall reliability was 0.97 for the entire instrument for this administration (table 3.1.3.1.3a).

Table 3.1.3.1.3a. November 2007 CCES High School scale mean, scale reliability, and instrument reliability

All Students Responding	<i>M</i>	α
Trustworthiness	3.5	0.89
Respect	3.6	0.90
Responsibility	3.3	0.84
Fairness	3.5	0.91
Caring	3.5	0.87
Citizenship	3.2	0.87
<i>N</i> =169		0.97

Response scale = 1-Almost Never 2-Rarely 3- Sometimes 4-Often 5-Almost Always

A MANOVA with administrations one (November 2007), two (February 2008), and three (March 2008) as independent variables (Time) and experimental group Trustworthiness,

Respect, Responsibility, Fairness, Caring, and Citizenship as the dependent variables demonstrated a statistically significant main effect [$F(12, 38) = 6.47, p = .000, \lambda = .108$, partial $\eta^2 = .671$, power = 1.00] of a drop then slight rise in students' perceptions of their environment over time.

Table 3.1.3.1.3b. Means and standard deviations of the High School CCES

	November 2007		February 2008		March 2008	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Trustworthiness	3.58	.31	3.34	.25	3.40	.21
Respect	3.76	.34	3.36	.22	3.32	.19
Responsibility	3.44	.17	3.13	.24	3.24	.22
Fairness	3.58	.13	3.14	.07	3.27	.07
Caring	3.58	.19	3.22	.18	3.32	.08
Citizenship	3.24	.20	2.95	.11	3.08	.11

Examination of experimental group results vs. control group results using MANOVA for the November 2007 administration indicated a statistically significant main effect [$F(6, 11) = 5.33, p = 0.008, \lambda = .256$, partial $\eta^2 = .74$, power = .92], where the experimental group subscale means were stronger than those of the control group (Table 3.1.3.1.3c).

Table 3.1.3.1.3c. November 2007 means by group and subscale

	Group	<i>M</i>	<i>SD</i>
Trustworthiness	1.00	3.57	.30
	2.00	3.43	.25
Respect	1.00	3.75	.33
	2.00	3.38	.22
Responsibility	1.00	3.44	.16
	2.00	3.24	.23
Fairness	1.00	3.57	.13
	2.00	3.30	.05
Caring	1.00	3.57	.18
	2.00	3.33	.14
Citizenship	1.00	3.24	.20
	2.00	3.12	.08

Experimental group is group 1 ($n = 155$); Control group is group 2 ($n = 54$)

Examination of experimental group results vs. control group results using MANOVA for the February 2008 administration indicated a statistically significant main effect [$F(6, 11) = 8.06, p = 0.002, \lambda = .185$, partial $\eta^2 = .82$, power = .99].

Table 3.1.3.1.3d. February 2008 means by group and subscale

	Group	<i>M</i>	<i>SD</i>
Trustworthiness	1.00	3.34	.24
	2.00	3.77	.43
Respect	1.00	3.35	.21
	2.00	3.77	.24
Responsibility	1.00	3.13	.24
	2.00	3.74	.51
Fairness	1.00	3.14	.07
	2.00	3.57	.30
Caring	1.00	3.22	.17
	2.00	3.68	.30
Citizenship	1.00	2.95	.11
	2.00	3.42	.31

Experimental group is group 1 ($n = 145$); Control group is group 2 ($n = 8$)

Examination of experimental group results vs. control group results using MANOVA for the March 2008 administration indicated a statistically significant main effect [$F(6, 11) = 5.33$, $p = 0.08$, $\lambda = .256$, partial $\eta^2 = .74$, power = .92].

Table 3.1.3.1.3e. March 2008 means by group and subscale

	Group	<i>M</i>	<i>SD</i>
Trustworthiness	1.00	3.40	.21
	2.00	3.10	.43
Respect	1.00	3.32	.18
	2.00	3.00	.40
Responsibility	1.00	3.24	.21
	2.00	3.07	.29
Fairness	1.00	3.26	.07
	2.00	2.95	.21
Caring	1.00	3.32	.08
	2.00	3.14	.34
Citizenship	1.00	3.07	.10
	2.00	2.73	.15

Experimental group is group 1 ($n = 110$); Control group is group 2 ($n = 21$)

3.1.4 To what extent did special education students internalize the six character traits?

Null Hypothesis (H_{01}) – There is no difference between classroom character-trait perceptions of experimental students when compared to control group students.

Population – Floresville ISD special education students (experimental group) and Stockdale ISD special education students (control group).

Sample – The sample was randomly selected from a list of teacher names in each school. The students of these teachers were then administered the instrument.

Instrument – Modified scales from a variety of pre-existing instruments and new items to measure students’ perceptions of six character-trait scales of: (1) trustworthiness, (2) respect, (3) responsibility, (4) fairness, (5) caring, and (6) citizenship in their classrooms was developed as the Classroom Community Environment Survey (CCES). The CCES was administered via the Web to reduce response errors. The CCES was customized for three levels of students: (1) elementary, with 18 items, three items per scale, (2) middle school, with 54 items, and a (3) high school version with 54 items.

3.1.4.1 Elementary-Level CCES for Special Education

The CCES was administered to a control group consisting of elementary school students and an experimental group of 115 students. Of the 115 experimental group students only six were identified as special education students. Of the 17 control group students none were identified as special education students. Therefore, the special education sample drawn from the larger random population was insufficient for conducting meaningful analyses.

3.1.4.2 Middle School Level CCESS for Special Education

The CCES was administered to a randomly selected experimental group of students of which only two were identifiable as special education. No control group results were available. Therefore, the special education sample drawn from the larger random population was insufficient for conducting meaningful analyses.

3.1.4.3 High School-Level CCES 2007 Baseline for Special Education

The CCES was administered to a control group consisting of high school students and an experimental group of 155 students. Of this population only 12 experimental group students were identifiable as special education and only two control group students were identifiable as special education. Therefore, the special education sample drawn from the larger random population was insufficient for conducting meaningful analyses.

3.2 OUTCOMES OF OBJECTIVE TWO

3.2.1 To what extent... (1) did students learn, (2) students’ attitudes change, and (3) were students impacted by service learning.

Study period April 2007

A new instrument, the 57-item *Service Learning Survey* (SLS), was developed with eight scales of: Learning (8 items), Autonomy (6 items), Challenge (3 items), Belonging (3 items), Empowerment (5 items), Responsibility (5 items), Attitude toward Community Activity (11 items), and Attitude toward Service (10 items). The SLS was content validated May 8, 2007, with 37 mixed-grade high school students, two of whom were special education classified. After content validation the 5-item response scale was

reduced to a 4-item response scale of *Never, Seldom, Often, and Always*. The semantic differential responses for the two attitude scales remained at five levels.

When subjected to Chronbach's Alpha, the overall instrument for the field test held up at 0.96. The individual scale alpha coefficient's ranged from 0.56 to 0.92, with the weakest being the scale of Challenge and the strongest the scale of student Attitudes toward Service (Table 3.2.1a).

Table 3.2.1a. Field test mean, SD, and alpha reliability coefficient for the SLS

Scale	<i>M</i>	<i>SD</i>	α
	Scale = 1 - 4		
Learning	2.95	4.6	0.81
Autonomy	2.96	3.4	0.78
Challenge	2.96	1.8	0.56
Belonging	2.58	2.2	0.77
Empowerment	2.95	3.2	0.78
Responsibility	2.8	3.4	0.83
	Scale = 1 - 5		
Attitude toward Community Activity	3.9	7.9	0.90
Attitude toward Service	4.0	7.9	0.92
Entire Instrument			0.96

N = 211

When subjected to factor analysis the scales of Learning, Autonomy, Challenge, Belonging, Empowerment, and Responsibility (analyzed together since they had the same four point response scale) mostly held to their a priori scale when the cut-off score was set at 0.40. Some Learning scale items crossed over presumably because the entire instrument was administered under the pretext of learning regarding service learning (Table 3.2.1b).

Table 3.2.1b. Service learning impact scales rotated component factor matrix for the SLS

Item	Learning	Autonomy	Challenge	Belonging	Empowerment	Responsibility
Learn6	.576					
Learn7	.434					
Learn8	.523					
Learn9	.669					
Learn10	.741					
Learn11			.447			
Learn12				.411		
Learn13					.427	
Autonomy14		.564				
Autonomy15		.700				
Autonomy16		.641				
Autonomy17		.610				
Autonomy18		.647				
Autonomy19		.413		.507		
Challenge20			.792			
Challenge21			.526			
Challenge22			.476			
Belong23				.616		
Belong24				.635		
Belong25				.677		
Empower26					.790	
Empower27					.589	
Empower28					.712	
Empower29					.802	
Empower30					.714	
Resp31						.745
Resp32						.775
Resp33						.640
Resp34						.570
Resp35						.651

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. *N* = 211. Factors below 0.40 are not shown.

When independently subjected to factor analysis the two attitude scales (analyzed independent of the impact scales since they have the same five point response scale) mostly held to their a priori scale expectations when the cut-off score was set at 0.50 (Table 3.2.1c). Given these analysis results in rotated component factor matrices Table 3.2.1b and 3.2.1c, the SLS demonstrates good validity.

Table 3.2.1c. Service learning attitude scales rotated component factor matrix for the SLS

Item	Attitude 2	Attitude 1
Attitude36		.644
Attitude37		.629
Attitude38		.630
Attitude39		.601
Attitude40		.787
Attitude41		.620
Attitude42		.800
Attitude43		.752
Attitude44	.579	
Attitude45		
Attitude46		.752
Attitude47	.541	
Attitude48	.627	
Attitude49	.643	
Attitude50	.718	
Attitude51	.737	
Attitude52	.755	
Attitude53	.861	
Attitude54	.712	
Attitude55	.546	
Attitude56	.807	

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. $N = 211$. Factors below 0.50 are not shown.

Following instrument validity and reliability analyses an independent-samples t -test was conducted to analyze if the means of the experimental group (students enrolled in the VALU Service Learning class) and the control group students (those not enrolled in the service learning class) differed significantly at the 95% confidence level. Levine's item-by-item test of heteroschedasticity was used to determine if the variance of the two groups' distributions differed significantly. At a 95% confidence level 11 items were found to have significant variances. Therefore, on the independent-samples t -test unequal-variance estimates were used rather than the equal-variance estimates for those 11 items.

The results of the Independent-Samples t -test indicated that **means of the experimental group and the control group differed significantly at the 95% confidence level** on 15 items (Table 3.2.1d).

Table 3.2.1d. SLS independent-samples *t*-test results

Scale	Item Number	<i>t</i>	<i>p</i>
Learning	6	4.64	.000
	8	2.19	.030
Autonomy	14	3.09	.002
	15	3.98	.000
	16	2.38	.018
Challenge	19	3.09	.002
	21	4.60	.000
	36	2.03	.044
Attitude toward Community Activity	43	3.03	.003
	44	2.49	.014
Attitude toward Service	47	2.47	.014
	50	2.23	.026
	53	2.64	.009
	54	2.71	.007
	55	2.39	.017

p < .05, *df* = 209

These data suggest that on the scales of Learning, Challenge, Belonging, Empowerment, Responsibility, and Attitude toward Community Activity there was very little difference between experimental group students and control group students. However, on the scales of Autonomy and Attitude toward Service, there is a difference overall between VALUE Service Learning class students and students not enrolled in the VALUE class.

In an effort to explore for any associations between the impact scales of Learning, Autonomy, Challenge, Belonging, Empowerment, and Responsibility, and the attitude scales, no statistically significant associations were found in simple correlation analysis (*r*) between how service was performed (item 4b; as part of a class, part of a school-sponsored activity, or as part of a non-school group) and the impact scales (Table 3.2.1e). And, the statistically significant correlations between the scales and enrollment in the VALUE Service Learning class were negative (Table 3.2.1e), suggesting that service learning course enrollment and how one does service, if s/he does it, does not have an impact on student attitudes toward service or other factors studied in this measure.

Table 3.2.1e. Associations between eight SLS scales and students enrolled in VALU classes

Scale	<i>r</i> for enrolled in VALU class	<i>r</i> for how service performed
Learning	-.07**	.20
Autonomy	-.15**	-.30
Challenge	-.09**	-.21
Belonging	-.04**	-.66
Empowerment	-.09**	-.27
Responsibility	-.04**	-.51
Attitude toward Community Activity	-.11**	.05
Attitude toward Service	-.11**	-.19

***p* < .01

3.2.2 Study period August 2007

The 56-item *Service Learning Survey* (SLS), with eight scales of: Learning (8 items), Autonomy (6 items), Challenge (3 items), Belonging (3 items), Empowerment (5 items),

and Responsibility (5 items), with a 4-item response scale of *Never, Seldom, Often, and Always*, and two semantic differential scales of Attitude toward Community Activity (11 items) and Attitude toward Service (10 items) was administered to 38 mixed-grade high school students in August 2007.

When subjected to Chronbach’s alpha, the overall instrument for the field test held up at 0.97. The individual scale alpha coefficient’s ranged from 0.60 to 0.94, with the weakest being the scale of Challenge and the strongest the scales being that of student Attitudes toward Service and Attitudes toward Community (Table 3.2.2a).

Table 3.2.2a. Field test mean, SD, and alpha reliability coefficient for the SLS

Scale	<i>M</i>	<i>SD</i>	α
	Scale = 1 - 4		
Learning	3.89	4.9	0.82
Autonomy	2.99	2.8	0.63
Challenge	2.75	3.8	0.60
Belonging	2.48	2.5	0.79
Empowerment	2.84	3.4	0.81
Responsibility	2.77	3.5	0.82
	Scale = 1 - 5		
Attitude toward Community Activity	3.68	9.6	0.94
Attitude toward Service	3.84	9.5	0.94
Entire Instrument			0.97

N = 38

Correlation analysis revealed good to strong correlations (Table 3.2.2b) between dependent variables, so Learning, Autonomy, Challenge, Belonging, and Empowerment were combined to form a new variable called Service Learning.

Table 3.2.2b. August 2007 dependent variable correlations

	Learning	Autonomy	Challenge	Belonging	Empowerment	Responsibility	Attitude
Learning	1	-.210	.596	.982	.151	.259	-.484
Autonomy		1	-.217	.786	.876	.175	-.099
Challenge			1	.434	.322	.986	.606
Belonging				1	.993	.577	-.454
Empowerment					1	.549	-.675
Responsibility						1	.114
Attitude							1

N = 21

A second correlation analysis between the new variable Service Learning and the previous Responsibility, and Attitude variables reveal weak correlations ($r < .26$).

3.2.3 Study period December 2007

When subjected to Chronbach’s alpha analysis, the overall instrument for the field test held up at 0.98. The individual scale alpha coefficient’s ranged from 0.67 to 0.96, with the weakest being the scale of Challenge and the strongest the scale of student Attitudes toward Community Activity (Table 3.2.3a).

Table 3.2.3a. Field test mean, SD, and alpha reliability coefficient for the SLS

Scale	<i>M</i>	<i>SD</i>	α
	Scale = 1 - 4		
Learning	3.13	4.6	0.85
Autonomy	3.06	2.7	0.79
Challenge	3.00	1.8	0.67
Belonging	2.74	2.3	0.84
Empowerment	2.98	3.7	0.89
Responsibility	2.81	3.8	0.87
	Scale = 1 - 5		
Attitude toward Community Activity	3.83	9.6	0.96
Attitude toward Service	4.00	8.4	0.95
Entire Instrument			0.98

N = 43

Correlation analysis revealed good to strong correlations (Table 3.2.3b) between dependent variables, so Learning, Autonomy, Challenge, Belonging, and Empowerment were combined to form a new variable called Service Learning. Responsibility and Attitude remained as dependent variables.

Table 3.3.3b. December 2007 dependent variable correlations

	Learning	Autonomy	Challenge	Belonging	Empowerment	Responsibility	Attitude
Learning	1	.239	.961	.693	.666	.515	-.352
Autonomy		1	.756	.982	.917(*)	.172	-.430
Challenge			1	.866	.933	.500	.000
Belonging				1	.988	.000	-.500
Empowerment					1	.528	-.608
Responsibility						1	.161
Attitude							1

* Correlation is significant at the 0.05 level (2-tailed).

A new correlation analysis was conducted on the new variable set revealing $r < .52$.

3.2.4 Study period August to April 2008 compared

In order to determine if service learning influenced students' perceptions, responsibility, and attitudes, a MANOVA was conducted with group (experimental vs. control) as the independent variable and August and December administrations of variables Service Learning Composite, Responsibility, and Attitude as the dependent variables. No significant main effect was found [$\lambda = .061$, $F(6, 3) = 7.67$, $p = .061$, $\eta^2 = .94$, power = .58].

In order to determine if service learning influenced students' perceptions, responsibility, and attitudes during the second round of evaluation, a MANOVA was conducted with group (experimental vs. control) as the independent variable and January and April administrations of variables Service Learning Composite, Responsibility, and Attitude as the dependent variables. No significant main effect was found [$\lambda = .119$, $F(6, 3) = 3.52$, $p = .154$, $\eta^2 = .88$, power = .33]. Likewise, MANOVA was used to investigate if students' perceptions changed over time with the August, December, January, and April results

as the dependent variable. No significant effect was found [$\lambda = .834$, $F(9, 34) = .294$, $p = .972$, $\eta^2 = .059$, power = .11].

3.3 OUTCOMES OF OBJECTIVE THREE

3.3.1 To what extent did elementary and middle school teachers' knowledge, awareness, and efficacy increase?

3.3.1.1 Building Developmental Assets – Study period December 2006 to April 2007

This measure was conducted in two phases. Phase I was a quantitative multiple time-series measure using a 10-item test of teachers' knowledge, awareness, and efficacy related to Building Developmental Assets (BDA) professional development, a mass-produced, pre-packaged character education teacher professional development program presented to 108 Primary and Elementary school teachers in January 2007. Phase II was a follow-up qualitative measure consisting of a focus group with a random sample of teachers ($N = 6$) who participated in the BDA professional development.

Phase I quantitative results

There was one statistically significant difference between the pre-test and post-test 1 at $p < .05$ using a paired-samples t -test (Table 3.3.1.1a) with no correlation between those scoring high on one test scoring high on another ($r = -0.004$, $p = 0.963$). The remaining paired-samples t -tests demonstrated no statistically significant differences between the test administrations.

Table 3.3.1.1a. Building Developmental Assets paired-sample t -test results

Pair	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>r</i>
Pre-test cognitive	42.6	15.3	118	-4.9	0.000	-0.004
Post-test 1 cognitive	54.3	21.2				
Pre-test efficacy	3.2	0.4	1	-3.5	0.180	1.000
Post-test 1 efficacy	3.5	0.3				
Post-test 1 cognitive	54.3	21.2	118	0.7	0.450	0.017
Post-test 2 cognitive	52.5	18.2				
Post-test 1 efficacy	3.5	0.3	1	-1.9	0.310	1.000
Post-test 2 efficacy	3.6	0.2				

$p < .05$ (two-tailed), $N = 119$

Using a one-sample t -test with a preset test value of 70 (a passing grade in many academic environments) teachers scored significantly below the target on each test administration (table 3.3.1b).

Table 3.3.1.1b. Building Developmental Assets one-sample t-test results

Test	M	SD	df	t	p
Pre-Test	41.3	15.3	171	-24.6	0.000
Post-Test 1	54.3	21.2	118	-8.1	0.000
Post-Test 2	52.6	19.2	142	-10.8	0.000

Test value = 70; $p < .05$

These quantitative results indicate that outside of an immediate post-professional development increase in character education knowledge at the simple recall-remember cognitive level (Anderson et al., 2001) teachers' overall knowledge, awareness, and efficacy regarding character education, strictly in terms of Building 40 Development Assets, had little lasting impact during the study period.

Phase II qualitative results

On May 22, 2007, three Primary teachers and three Elementary teachers were randomly selected ($N = 6$) to participate in a one-hour long focus group to speculate on reasons behind selected quantitative test results. The selected results were those that appeared as outliers or extreme cases when descriptive test scores from each administration of the Building Developmental Assets (BDA) 10-item measure of knowledge, awareness, and efficacy were presented (Figures 3.3.1.1a & b). Selected test items discussed included items:

3. *Support* can be practiced by which of the following?
4. *Social Competencies* can be practiced by which of the following?
7. The primary developmental asset schools can most directly impact is/are:
8. Which one of the following is a "new attitude" according to the Search Institute?
9. How competent do you feel as a Developmental Assets teacher?
10. Overall, how effective do you feel you are at teaching "character"?

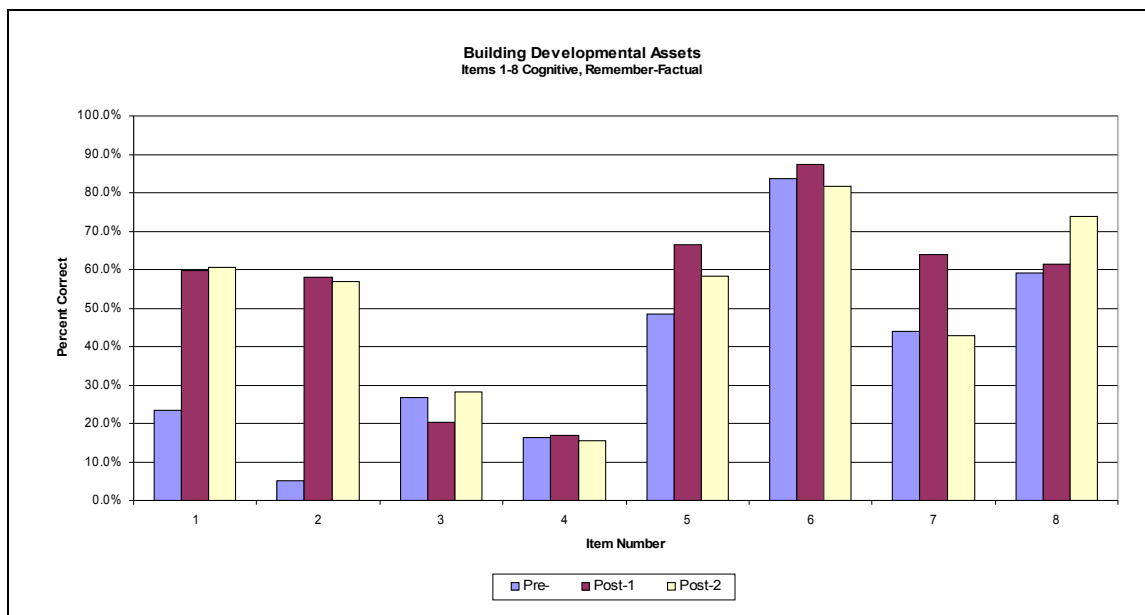


Figure 3.3.1.1a. BDA mean cognitive results

Regarding item 3, the low overall mean was attributed by the focus group as material not learned during the Building Developmental Assets (BDA) professional development that was held in January 2007 (after the pre-test).

The group indicated that the consistently low mean on item 4 could be attributed to the lack of context clues in the response items (there were four response choices—one correct and three distracters). In other words, they could not guess the correct answer. The group also noted that a lot of material was presented during the training (“...a lot to take in...”) and they were not certain what they were supposed to be learning at that point.

Regarding item 7, the group consensus was that they did not learn the material covered in the item having to do with achievement motivation as the primary developmental asset schools could most directly impact.

The consistent mean rise in knowledge reflected in item 8 was attributed by the focus group to their previous knowledge rather than that acquired in the BDA professional development. One teacher noted that students as resources “reflects student-centered learning.”

Efficacy items 9 and 10, ranked on a Likert-like scale of 1 to 5, were presented and teachers indicated the consistent rise in the mean could be attributed to more exposure to character education and the simple fact that over time and throughout the school year teachers feel more competent and effective.

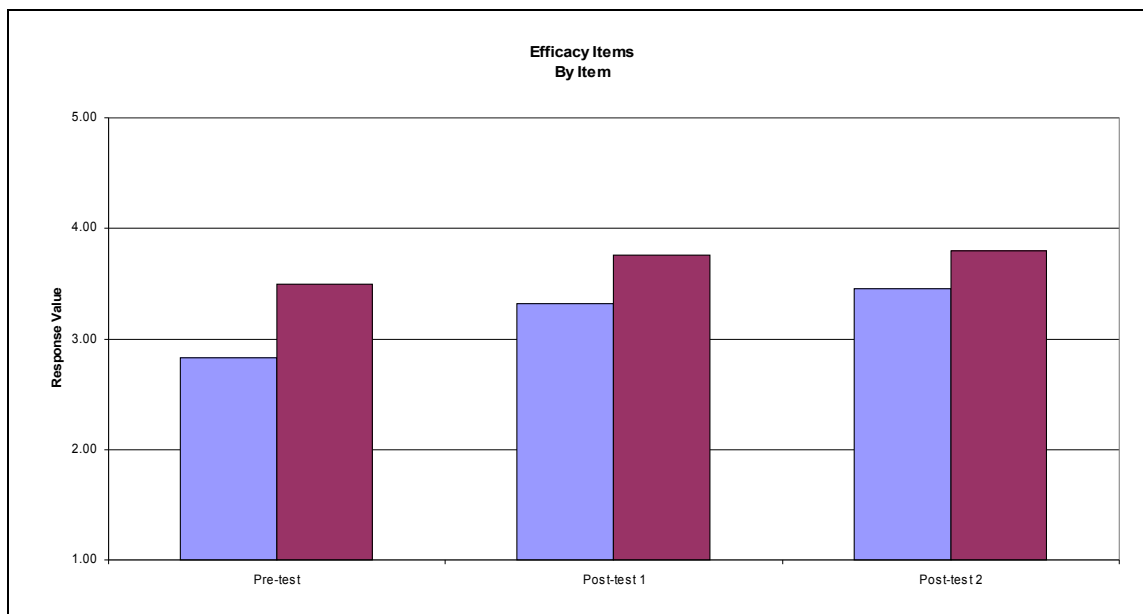


Figure 3.3.1.1b. BDA mean efficacy results

Powerful Teaching – Study period December 2006 to April 2007

This measure was conducted in two phases. Phase I was a quantitative multiple time-series measure using a multi-response, 7-item test of teachers’ knowledge, awareness, and efficacy related to Powerful Teaching (PT) professional development, a mass-produced, pre-packaged character education teacher professional development program presented to 60 Middle School teachers. Phase II was a follow-up qualitative measure consisting of a focus group with a random sample of teachers ($N = 4$) who participated in the PT professional development.

Phase I quantitative results

In a series of paired-sample t -tests, the difference in means from the pre-test to post-test 1 on the measure of teacher knowledge was statistically significant ($p < .05$). Likewise, the difference in means from the pre-test to post-test 1 on the measure of efficacy was statistically significant (Table 3.3.1.1c). The remaining means demonstrated no statistically significant differences and there were no correlations between those scoring high on one test scoring high on another.

Table 3.3.1.1c. Powerful Teaching paired-sample t -test results

Pair	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>	<i>r</i>
Pre-test cognitive	53.7	9.67				
Post-test 1 cognitive	60.6	15.1	47	-2.9	0.006	0.142
Pre-test efficacy	2.9	1.2				
Post-test 1 efficacy	3.7	0.9	47	-3.3	0.002	-0.65
Post-test 1 cognitive	60.6	15.1	25	-1.1	0.278	0.18
Post-test 2 cognitive	63.7	14.2				
Post-test 1 efficacy	3.7	0.9	25	0.2	0.882	0.11
Post-test 2 efficacy	3.6	0.9				

$p < .05$ (two-tailed), $N = 48$

Using a one-sample t -test with a preset test value of 70 (a passing grade in many academic environments) teachers scored significantly below the target on the pre-test and post-test 1 (Table 3.3.1.1d). However, on post-test 2 the data do not demonstrate a significantly low score.

Table 3.3.1.1d. Powerful Teaching one-sample t -test results with test value of 70

Test	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Pre-Test	53.9	9.3	60	-13.4	0.000
Post-Test 1	60.6	15.1	47	-4.3	0.000
Post-Test 2	63.7	14.5	25	-2.2	0.035

Test value = 70; $p < .05$

These results indicate there is a chance that teachers’ gained knowledge related to Powerful Teaching as they practiced and became more familiar with the methods.

Phase II qualitative results

On May 22, 2007, four Middle School teachers were randomly selected to participate in a 45-minute long focus group to speculate on reasons behind selected quantitative test results. The selected results were those that appeared as outliers or extreme cases when descriptive test scores from each administration of the Powerful Teaching (PT) 7-item measure of knowledge, awareness, and efficacy were presented (Figures 3.3.1.1c & d). Overall test results and the following selected test items were discussed:

2. List as many strategies as you can for infusing Developmental Assets in your classroom/content area:

5. There are 8 pro-social asset categories with 4 of them representing external structures, relationships, and activities that create a positive environment for young people. The other 4 categories reflect internal values, skills, and beliefs young people need to fully engage and function.

Please read over the following list and select whether the category focuses on external or internal assets.

Asset Category	External Assets	Internal Assets
Boundaries & expectations	___	___
Commitment to learning	___	___
Constructive use of time	___	___
Empowerment	___	___
Positive identity	___	___
Positive values	___	___
Social competencies	___	___
Support	___	___

6. We know that asset building in classrooms and schools can occur in three key areas. For each area, offer a tangible example of what you are doing/could do to focus more intentionally on asset building:

- a. Relationships:
- b. The classroom environment:
- c. Your classroom curriculum:

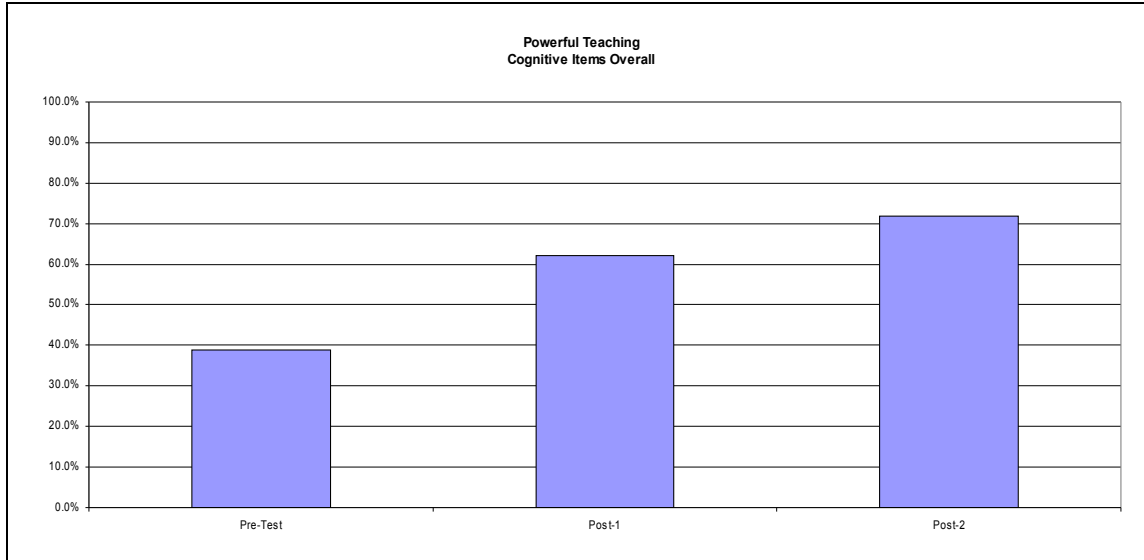


Figure 3.3.1.1c. Powerful Teaching overall mean test scores, three administrations (N = 60)

Regarding the overall rising trend (Figure 3.3.1.1c) in mean scores of 60 teachers, the focus group teachers attributed this to teachers simply being more aware of character education after the professional development (post-test 1). They also agreed that they had some awareness of character education after being involved in a Middle School principal-initiated book study prior to the grant period and that after the Power Teaching professional development much of the book and the PT program became more evident and concrete.

When the focus group was asked to speculate on the sharp rise then drop in the mean score for item 2 (Figure 3.3.1.1d) there was consensus that during post-test 1 “could do’ ideas were fresh on our minds.” The drop indicated in post-test 2 was likely due to teachers’ focus on the state-mandated achievement test (Texas Assessment of Knowledge and Skills).

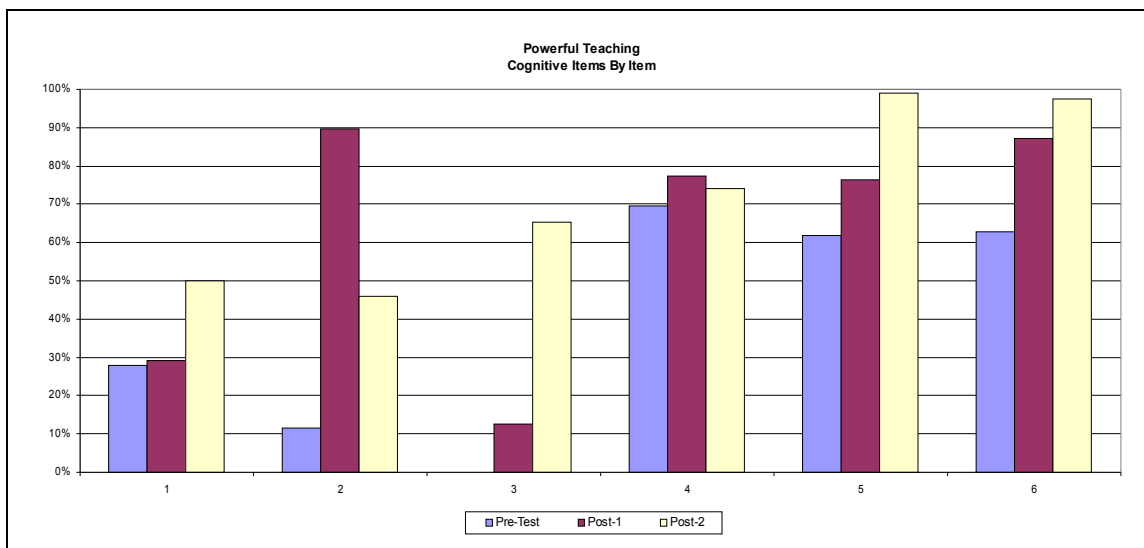


Figure 3.3.1.1d. Powerful Teaching mean scores by item

When queried regarding the consistent rise in mean scores for items 5 and 6 (Figure 3.3.1.1d) the group admitted to using their Powerful Teaching workbook to aid them in recalling facts. They noted that most of the teachers had figured out by post-test 2 that these items could be answered from the book.

3.3.1.2 Smart and Good Schools - Study period August 2007 – April 2008

Building Developmental Assets and Powerful Teaching measures were dropped for the 2007-2008 school year and replaced with the Smart and Good Schools (S&GS) measure due to changes in the professional development program adopted by Floresville ISD.

The Smart and Good Schools instrument was administered in October 2007, January 2008, and March 2008 to the experimental group and the control group (note: no control group data were collected for March 2008 Middle School teachers, no Elementary-level teacher data was provided at all for March 2008).

Middle School Results

MANOVA results of group by score percentage and efficacy scale mean demonstrated no statistically significant main effect [$\lambda = .956$, $F(4, 67) = .763$, $p = .553$, $\eta^2 = .044$, power = .233]. MANOVA results for the experimental group over time demonstrated a statistically significant main effect [$\lambda = .625$, $F(4, 358) = 23.75$, $p = .000$, $\eta^2 = .210$, power = 1.0] whereby the mean test score of each administration dropped over time [October $M = 51$, $SD = 16.0$, $n = 61$; January $M = 29$, $SD = 11.3$, $n = 79$; March $M = 29$, $SD = 12.2$, $n = 42$] while efficacy scale mean ($M = 3.5$) held steady on a 1 to 5 scale (Table 3.3.1.2a).

Table 3.3.1.2a. Middle school mean scores over time

	N	Score Mean	SD	Efficacy Mean	SD
October 2007	61	51	16.0	3.5	.69
January 2008	79	29	11.3	3.6	.68
March 2008	42	29	12.2	3.5	.67

Primary School Results

MANOVA results of group by score percentage and efficacy scale mean demonstrated a statistically significant main effect [$\lambda = .863$, $F(4, 70) = 2.78$, $p = .033$, $\eta^2 = .137$, power = .735]. A follow-up ANOVA revealed the October administration's between groups scores $F(1, 103) = 4.35$ was significant at $p = .039$, where the experimental group test score percentage $M = 41$, $SD = 15.4$, $n = 92$ and the control group test score percentage $M = 51$, $SD = 17.2$, $n = 12$). Likewise, the January administration demonstrated between groups scores as $F(1, 76) = 4.63$ was significant at $p = .035$, where the experimental group test score percentage $M = 26$, $SD = 12.2$, $n = 66$ and the control group test score percentage $M = 34$, $SD = 10.3$, $n = 14$). Efficacy perceptions were not significantly different.

MANOVA results for the experimental group over time (October and January) demonstrated a statistically significant main effect [$\lambda = .796$, $F(2, 152) = 19.5$, $p = .000$, $\eta^2 = .204$, power = 1.0] whereby the mean test score of each administration dropped over time [October $M = 41$, $SD = 15.4$, $n = 92$; January $M = 26$, $SD = 11.3$, $n = 66$] while the efficacy scale mean ($M = 3.3$) held steady on a 1 to 5 scale (Table 3.3.1.2b).

Table 3.3.1.2b. Experimental group over time

	N	Score Mean	SD	Efficacy Mean	SD
October 2007	92	41	15.4	3.3	.67
January 2008	66	26	12.2	3.3	.68

3.3.2. To what extent did administrators' knowledge, awareness, and efficacy increase?

3.3.2.1 Leadership Academy in Character Education (LACE) Test of Knowledge and Efficacy – Study period October 2007

The Leadership Academy in Character Education (LACE) Test of Knowledge and Efficacy was developed specifically to assess school administrators' knowledge and efficacy related to character education. The 13-item instrument includes 11 knowledge items and two efficacy items (perceived competence and perceived effectiveness). The LACE Test was administered for the first time in October 2007 to 22 district administrators. Chronbach's alpha was low at 0.42, with inter-item correlation at 0.50 and below. The mean Knowledge score was 74, $SD = 16.2$ on a scale of 0 -100 and the Efficacy $M = 3.5$, $SD = .95$ on a scale of 1 to 5, with 5 being the strongest. In addition to general descriptive analyses, simple linear correlation analysis found no statistically significant association between administrator knowledge and efficacy ($r = -0.19$, $p = 0.401$).

3.3.2.2 Leadership Academy in Character Education (LACE) Test of Knowledge and Efficacy – Study period February 2008

The LACE Test was administered again in February 2008 ($N = 12$). Chronbach's alpha was low at 0.16, with inter-item correlation at 0.45 and below. The mean Knowledge score was 62, $SD = 14.2$ on a scale of 0 -100 and the Efficacy $M = 3.8$, $SD = .53$ on a scale of 1 to 5, with 5 being the strongest. In addition to general descriptive analyses, simple linear correlation analysis found no statistically significant association between administrator knowledge and efficacy ($r = 0.27$, $p = 0.35$).

3.3.2.3 Leadership Academy in Character Education (LACE) Test of Knowledge and Efficacy – Study period October 2007 and February 2008

When a One-Sample t -test with a 75 cut-off score was conducted on the October 2007 results the mean percent was not significantly lower than the cut-off score ($t = -.34$, $df = 21$, $p = .74$). The results of the One-Sample t -test on the February 2008 data revealed a significantly lower mean percent than the cut-off score ($t = -3.2$, $df = 11$, $p = .009$). A Paired-Samples t -test found the knowledge scores on the second test ($M = 62$) were significantly lower than those on the first test ($M = 74$) ($t = 3.23$, $df = 11$, $p = .008$). There was no association between variables ($r = .11$, $p = .73$). A Paired-Samples t -test found the efficacy levels indicated on the second test ($M = 3.8$) were not significantly different than those on the first test ($M = 3.5$) ($t = -1.86$, $df = 23$, $p = .08$). Likewise, there was no association between variables ($r = .08$, $p = .71$).

3.3.3 To what extent did parents and community leaders' knowledge and efficacy increase?

3.3.3.1 Study period December 2006 to May 2007

This measure consisted of a single-group prettest-posttest design and a quasi-experimental nonequivalent-groups posttest-only design. The single-group pret-test—a 10-item quiz of knowledge, awareness, and efficacy—was administered at the beginning of a meeting of parents and community leaders ($N = 22$). A post-test was administered at the end of the information session. Short-term duration knowledge and awareness was significant ($p = .001$) and somewhat correlated ($r = 0.48$). Short-term perceptions of efficacy demonstrated significance ($p = .005$) and correlation ($r = .61$) (table 3.3.3.1a).

Table 3.3.3.1a. Single-group prettest-posttest paired-sample t -test results

Test	Mean	SD	α	r	t	p
Pre-test cognitive	44.3	22.4	.66	0.48	-3.9	0.001
Post-test cognitive	63.1	20.9	.64			
Pre-test efficacy	3.48	0.85	.66	0.61	-3.0	0.005
Post-test efficacy	3.83	0.81	.70			

$p < 0.05$, $N = 22$

The quasi-experimental nonequivalent-groups posttest-only results used the post-test from the above experimental group ($n = 22$) and a single test from a control group ($n = 11$). The results of an independent-samples t -test indicate a statistically significant difference between experimental group scores and control group scores on knowledge ($p = .000$) and efficacy ($p = .000$) items.

Table 3.3.3.1b. Quasi-experimental nonequivalent-groups post-test independent-sample t -test results

Test	Mean	SD	t	df	p
Experimental group cognitive	63.0	20.9	5.38	31	0.000
Control group cognitive	25.0	14.7			
Experimental group efficacy	3.6	1.1	5.77	31	0.000
Control group efficacy	1.77	0.7			

Equal variance assumed based on Levene's Test. $p < 0.05$.

Finally, a one-sample t -test was conducted using a cut-off score of 70 (a passing grade in many academic environments). Community leaders and parents (experimental group, $n = 22$) scored significantly below the target on the pre-test ($t = -5.4$, $df = 21$, $p = .000$) and the control group ($n = 11$) also scored significantly below the target ($t = -10.1$, $df = 10$, $p = .000$).

Table 3.3.3.1c. One-sample t -test results

Test	M	SD	df	t	p
Experimental Group Pre-Test	44.3	22.4	21	-5.3	0.000
Experimental Group Post-Test	63.1	20.9	21	-1.5	0.136
Control Group	25.0	14.8	10	-10.1	0.000

Test value = 70; $p < .05$ (two-tailed)

These results should be viewed with the limits in mind that the community leaders/parents (experimental group) post-test was conducted immediately after the information meeting. It should also be noted that the control group mean (25) and the experimental group pre-test mean (44) indicate that some community leaders/parents came in with more knowledge and awareness, which is possibly why they participated in the meeting, simply because they are more aware and active parents/community leaders.

3.3.3.2 Study period March 2008 to May 2008

A new instrument was developed to measure a newly-implemented community professional development program called the Common Sense Parenting and Character Education class. The new instrument consisted of an 8-item quiz of knowledge and efficacy. It was administered in March 2008 to three course participants (experimental group) and 10 non-participants (control group). The results are presented in the table below.

Table 3.3.3.2 March 2008 CSPCE test results

	Knowledge		Efficacy	
	M	SD	M	SD
Experimental Group	83	29.4	3.0	.00
Control Group	57	12	2.2	.67

N = 13

Results of an independent-samples *t*-test indicated a significant difference between Knowledge mean percent scores ($t = 2.4$, $df = 11$, $p = .034$) and between Efficacy scale scores ($t = 3.1$, $df = 24$, $p = .005$).

3.3.4 To what extent did high school teachers' knowledge, awareness, and efficacy increase?

3.3.4.1 Study period December 2006 to April 2007

This measure was conducted in two phases. Phase I was a quantitative multiple time-series measure using an 11-item test of teachers' knowledge, awareness, and efficacy related to Smart and Good High Schools (SGHS) professional development, a mass-produced, pre-packaged character education teacher professional development program presented to 86 High School teachers. Phase II was a follow-up qualitative measure consisting of a focus group with a convenience sample of teachers ($N = 6$) who participated in the SGHS professional development.

Phase I quantitative results

In a series of paired-sample *t*-tests, the difference in means demonstrated no statistically significant differences (Table 3.3.4.1a). There was a correlation demonstrated between those scoring higher on the pre-test also scoring high on post-test 1 ($r = .70$).

Table 3.3.3.1a. Smart and Good High Schools paired-sample *t*-test results

Pair	M	SD	df	<i>t</i>	<i>p</i>	<i>r</i>
Pre-test cognitive	49.1	15.7				
Post-test 1 cognitive	55.3	19.1	49	-1.7	0.900	0.70
Pre-test efficacy	3.3	0.3				
Post-test 1 efficacy	3.4	0.2	1	-1.3	0.421	0.00
Post-test 1 cognitive	55.3	19.1				
Post-test 2 cognitive	47.3	20.1	49	2.3	0.029	0.18
Post-test 1 efficacy	3.7	0.9				
Post-test 2 efficacy	3.6	0.9	25	0.2	0.882	0.21

$p < .05$ (two-tailed)

Using a one-sample *t*-test with a preset test value of 70 (a passing grade in many academic environments) teachers scored significantly below the target on all three tests

(Table 3.3.4.1b). However, it should be considered that only the pre-test had a normal distribution.

Table 3.3.4.1b. One-sample *t*-test results with test value of 70

Test	<i>M</i>	<i>SD</i>	<i>df</i>	<i>t</i>	<i>p</i>
Pre-Test	47.9	16.5	94	-13.0	0.000
Post-Test 1	55.3	19.1	49	-5.4	0.000
Post-Test 2	47.8	19.2	59	-8.9	0.000

Test value = 70; *p* < .05 (two-tailed)

These results indicate that there is a chance that teachers' gained little knowledge related to Smart and Good High Schools training since their pre-test mean scores and their time-delayed post-test 2 mean scores are nearly equivalent.

Phase II qualitative results

On May 22, 2007, six High School teachers were selected to participate in a one-hour long focus group to speculate on reasons behind selected quantitative test results. The selected results were those that appeared as outliers or extreme cases when descriptive test scores from each administration of the Smart and Good High Schools 11-item measure of knowledge, awareness, and efficacy were presented (Figure 3.3.3a). Overall test results and the following selected test items were discussed:

1. Identify the six principles by which an ethical learning community is created.
2. *Performance* and *performance character* are the same thing.
3. *Task orientation* is when a student seeks to:
6. A Professional Ethical Learning Community (PELC) consists of:
7. List specific things you do to promote an ethical learning community in your classroom.
8. Which of the following are considered promising practices for developing diligent and capable performing students?
9. Which of the following are considered promising practices for developing an ethical learning community?

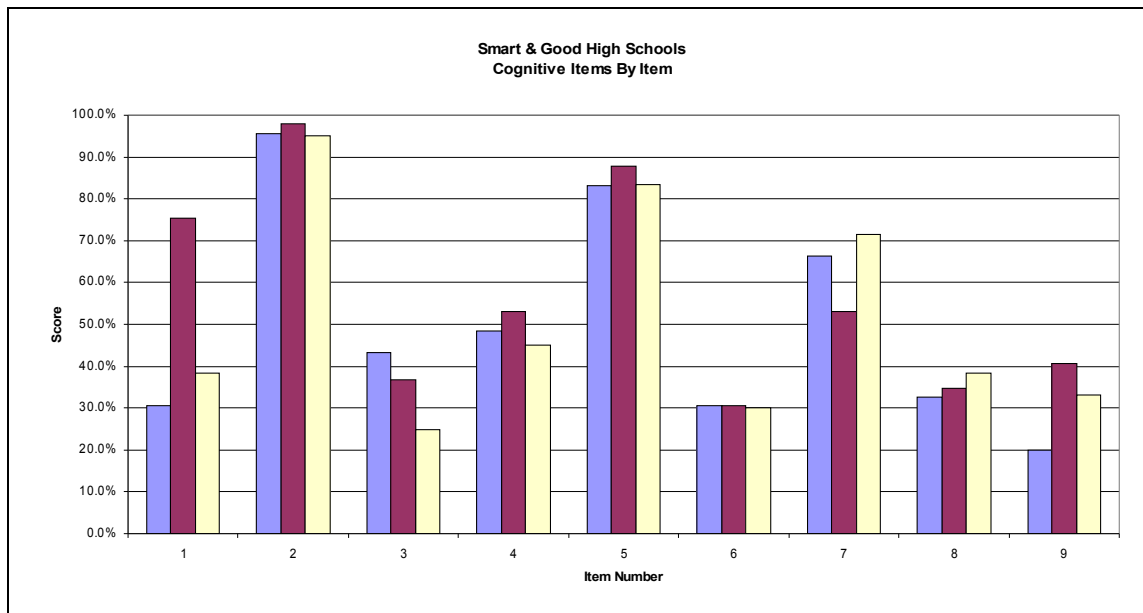


Figure 3.3.4.1a. Smart and Good High Schools cognitive item results

Regarding item 1 with a spiking mean shortly after the SGHS professional development and then a sharp drop during the second post-test, the focus group agreed that they were confused about the six pillars presented in the professional development and that they did not yet have sufficient information to select a more accurate response.

Item 2 had a particularly high mean for each test administration because it was the only alternative-response (true/false) item and the group speculated that the correct response was easy to guess since it was 50-50.

The low and consistently falling mean for item 3 the group suggested was related to what one teacher coined the “apathy factor.” Teachers admitted to just “getting the test done but no worrying too much about it.”

The consistently low mean for item 6 was attributed to uncertainty regarding the word “community” in the item. Teachers, the focus group suggested, may have thought community was related to the city of Floresville, not the school community as presented in the professional development.

The mean score recovery in item 7 was attributed to experience in recognizing ethical learning community and being able to practice teaching the concept in class.

Regarding overall low means for items 8 and 9, the teachers suggested all of the response items were “good answers.” In other words, the teachers could not discriminate between promising practices for developing *capable students* and for developing an *ethical learning community*, two distinctly different concepts, as presented in the SGHS professional development.

3.3.4.2 Study period August 2007 to April 2008

The Smart and Good Schools instrument was administered in October 2007, January 2008, and March 2008 to the experimental group and the control group (note: no control group data were collected for March 2008). Instrument reliability was low at $\alpha = .006$ October, $\alpha = .028$ January, $\alpha = .013$ March. MANOVA results of group by score percentage and efficacy scale mean demonstrated no statistically significant main effect [$\lambda = .923$, $F(4, 67) = 1.41$, $p = .242$, $\eta^2 = .077$, power = .414]. MANOVA results for the experimental group over time demonstrated a statistically significant main effect [$\lambda = .675$, $F(4, 406) = 22.01$, $p = .000$, $\eta^2 = .178$, power = 1.0] whereby the mean test score of each administration dropped over time [October $M = 50$, $SD = 17.0$, $n = 81$; January $M = 31$, $SD = 14.5$, $n = 63$; March $M = 29$, $SD = 10.58$, $n = 65$) while efficacy scale means held steady on a 1 to 5 scale (Table 3.3.4.2).

Table 3.3.4.2. Smart and Good Schools mean scores over time

	N	Score Mean	SD	Efficacy Mean	SD
October 2007	81	50	17.0	3.4	.72
January 2008	63	31	14.5	3.4	.81
March 2008	65	29	10.6	3.6	.59

3.3.5 To what extent did students perceive the transitional environment?

3.3.5.1 Study period May 2007

A new instrument, the High School Transition Survey (HSTS), was developed to measure the extent to which student perceive the transitional environment related to the Link Crew Program aimed at supporting students transitioning from 8th grade to high school.

The 18-item instrument consisted of a *Connection to School* scale and a *Self-Efficacy/Attribution* scale derived from modifications of previously developed instruments (Resnick et al, 1997; Skinner, Wellborn, & Connell, 1990; Stoltz, 2005). The instrument was content validated May 8, 2007, with 40 eighth grade students, 5 of which were special education classified.

The instrument was transformed into a Web-based instrument and administered on May 16, 2007. Only 21 responses were received which is insufficient to use for evaluation.

3.3.5.2 Study period October 2007

The HSTS was administered in October 2007 to 270 Floresville ISD (experimental group) incoming freshman and to 100 control group incoming freshman ($N = 370$). Reliability analysis was conducted and Chronbach's alpha of the entire instrument was 0.91. The alpha coefficient was 0.87 for the scale of *Connection to School* and 0.88 for the scale of *Self-Efficacy/Attribution*.

In order to examine the relationship between the scale of *Connection to School* and the scale of *Self-Efficacy/Attribution* a correlation analysis was conducted on the entire data set ($N = 370$). Results (Table 3.3.5.2) a moderate relationship among the variables, suggesting a composite high school transition variable is more suitable for later analyses.

Table 3.3.5.2. Correlation coefficients for transition variables

Variables	Connection to School	Self-Efficacy/Attribution
Connection to School	1.00	.573
Self-Efficacy/Attribution	.573	1.00
Mean	3.01	3.15
SD	.20	.33

$p < 0.01 = 0.14$ (2-tailed), $N = 370$

In order to explore differences between the experimental incoming freshman ($n=270$) and the control group ($n=100$), an Independent-Samples t -test was conducted. The test established that there was no statistically significant difference at the 95% confidence level between students who participated in the Link Crew program and students who did not ($t = .562$, $df = 34$, $p = 0.578$) where the experimental $M = 3.1$, $SD = .29$, and the control $M = 3.1$, $SD = .24$.

3.3.5.3 Study period May 2008

The High School Transition Survey (HSTS) was administered again in May 2008 to 228 Floresville ISD (experimental group) outgoing 8th grade students and to 83 control group outgoing 8th grade students ($N = 311$). Reliability analysis was conducted and Chronbach's alpha for the entire instrument was strong at 0.91. The alpha coefficient was 0.86 for the scale of *Connection to School* and 0.89 for the scale of *Self-Efficacy/Attribution*.

In order to examine the relationship between the scale of *Connection to School* and the scale of *Self-Efficacy/Attribution* a correlation analysis was conducted on the entire data set ($N = 311$). Results (Table 3.3.5.3) indicated weak relationships among the variables suggesting a composite high school transition variable is more suitable for later analyses.

Table 3.3.5.3 Correlation coefficients for transition variables

Variables	Connection to School	Self-Efficacy/Attribution
Connection to School	1.00	-.078
Self-Efficacy/Attribution	-.078	1.00
Mean	2.98	3.19
SD	.17	.30

$p < 0.01 = 0.842$ (2-tailed), $N = 311$

4. DISCUSSION

4.1 PROBLEMS

Program evaluation on a program this complex, with 29 evaluation measures in all, is difficult to roll out, coordinate, and maintain a viable number of responses. The first problem we experienced was related simply to diffusion of treatment due to the small size of the school district. Early in the grant it became impossible for the Floresville Independent School District (FISD) control group to NOT be exposed to character education initiatives.

A phenomenon that can best be termed “survey fatigue” was also quickly apparent with the interrupted-time series and randomized time series evaluation design. Teachers and students alike have the feeling of being “surveyed to death.” For instance, an incoming freshman might respond to the High School Transition Survey, the Service Learning Survey three times, and the Classroom Community Environments Survey twice for a total of six instruments in one school year. Teachers administering these instruments and responding to instruments and tests designed for the teachers themselves understandably find weariness in participating in yet another survey. Focus group results tell the tale of teachers checking off any random response just so they can get their instruments completed and get on to the next thing.

Other barriers include high teacher turnover (over 17%), the lack of instrument completion by control group teachers and students outside of our school district, attempts to establish consistently meaningful instruments in an ever changing and evolving program, and attempting to yield worthwhile evaluation results with very small numbers of participants, especially when randomizing the population from those who have given active consent.

4.2 LESSONS LEARNED

A primary lesson is that communication is essential throughout all program aspects. Survey fatigue, it was learned, can be combated by reducing instruments or combining them to reduce the number of times any given participant is asked to respond. Likewise, constant vigilance regarding IRB approval/renewal, compiling data, cleaning data, and analyzing data are important when data from 12 instruments administered multiple times and a variety of qualitative methods are used.

5. REFERENCES

- Anderson, L. W. et al. (Eds). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman: New York.
- Dorman, J. P. (2003). Cross-national validation of the What Is Happening In This Class? (WIHIC) questionnaire using confirmatory factor analysis. *Learning Environments Research: An International Journal*, 6(3), 231-245.
- Fraser, B. J. (1986). *Classroom environment*. London: Croom Helm, Ltd.
- Garson, D. (2001). *Factor analysis*. Retrieved July 29, 2006, from <http://www2.chass.ncsu.edu/garson/pa765/factor.htm>
- Kim, J., & Mueller, C. W. (1978). *Introduction to factor analysis: What it is and how to do it*. London: Sage Publications.
- Resnick, M. D. et al. (1997). Protecting adolescents from harm: Findings from the National Longitudinal Study on Adolescent Health. *The Journal of the American Medical Association*, 278, 823-832.
- Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: A process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology*, 82, 22-32.
- Stoltz, A. D. (2005). *The relationship between peer mentoring program participation and successful transition to high school*. Doctoral dissertation, University of California – Davis.
- Texas Education Agency. (2007). *Academic excellence indicator system 2006-2007 academic performance*. Austin, TX. Retrieved August 31, 2008, from <http://www.tea.state.tx.us/perfreport/aeis/>
- U.S. Census Bureau. (2000). Wilson County. Retrieved September 5, 2008, from <http://www.census.gov/>
- Walker, S. L. (2003). *Development and validation of an instrument for assessing distance education learning environments in higher education: The distance education learning environments survey (DELES)*. Doctoral dissertation, Curtin University of Technology, Perth, Western Australia, Australia. Retrieved, November 16, 2007, from <http://adt.curtin.edu.au/theses/available/adt-WCU20040211.095711/>

6. APPENDICES

APPENDIX I

Classroom Community Environment Survey (CCES)

Teacher Version 2.0

Teacher Instructions

Thinking of your students in general, please respond to the items below the best you can. There are no right or wrong responses.

Please select one response for each item below.					
1-Almost Never 2-Rarely 3- Sometimes 4-Often 5-Almost Always					
In my classes, students... [Trustworthiness]					
1. Tell the truth.	1	2	3	4	5
2. Believe honesty is the best policy.	1	2	3	4	5
3. Are trustworthy.	1	2	3	4	5
4. Avoid cheating on school work.	1	2	3	4	5
5. Keep promises they make to other students.	1	2	3	4	5
6. Are loyal to other students.	1	2	3	4	5
7. Avoid lying to avoid consequences.	1	2	3	4	5
8. Are honest.	1	2	3	4	5
9. Are generally truthful.	1	2	3	4	5
In my classes, students... [Respect]					
10. Treat classmates with respect.	1	2	3	4	5
11. Respect the personal property of others.	1	2	3	4	5
12. Show respect for school property.	1	2	3	4	5
13. Behave respectfully toward school staff.	1	2	3	4	5
14. Show respect toward teachers.	1	2	3	4	5
15. Show good sportsmanship.	1	2	3	4	5
16. Are respectful toward their schoolmates.	1	2	3	4	5
17. Refrain from put-downs.	1	2	3	4	5
18. Do not bully one another.	1	2	3	4	5
In my classes, students... [Responsibility]					
19. Break promises.	1	2	3	4	5
20. Are late to class.	1	2	3	4	5
21. Feel responsible in helping each other.	1	2	3	4	5
22. Get their school work done on time.	1	2	3	4	5
23. Take responsibility for their actions.	1	2	3	4	5
24. Pursue excellence in their school work.	1	2	3	4	5
25. Exercise self-restraint.	1	2	3	4	5
26. Report things that are wrong.	1	2	3	4	5
27. Are generally moral.	1	2	3	4	5
In my classes, students... [Fairness]					
28. Are concerned with equality.	1	2	3	4	5
29. Tend to be impartial.	1	2	3	4	5
30. Are open to one another.	1	2	3	4	5
31. Are equitable.	1	2	3	4	5
32. Are unbiased to one another.	1	2	3	4	5

33. Consider fairness in working with each other.	1	2	3	4	5
34. Pay attention to things being equal for one another.	1	2	3	4	5
35. Are evenhanded in their classroom interactions.	1	2	3	4	5
36. Are open to other students' ideas.	1	2	3	4	5
37. Practice parity.	1	2	3	4	5
In my classes, students... [Caring]					
38. Make an effort to include those who are different.	1	2	3	4	5
39. Try to comfort peers who have experienced sadness.	1	2	3	4	5
40. Help each other, even if they are not friends.	1	2	3	4	5
41. Work well together.	1	2	3	4	5
42. Help new students feel accepted.	1	2	3	4	5
43. Are willing to forgive each other.	1	2	3	4	5
44. Are patient with each other.	1	2	3	4	5
45. Listen to each other in class discussions.	1	2	3	4	5
46. Share what they have with others.	1	2	3	4	5
In my classes, students... [Citizenship]					
47. Do the right thing even if it's not in their best interest.	1	2	3	4	5
48. Resolve conflicts without fighting.	1	2	3	4	5
49. Try to get their friends to follow the rules.	1	2	3	4	5
50. Take an active role in solving classroom problems.	1	2	3	4	5
51. Consider each other's opinions.	1	2	3	4	5
52. Work together to bring about changes in the classroom community.	1	2	3	4	5
53. Obey classroom rules.	1	2	3	4	5
54. Tend to give more than they take.	1	2	3	4	5
55. Have the opportunity to have a voice in classroom governance.	1	2	3	4	5

APPENDIX II

Classroom Community Environment Survey (CCES) Primary/Elementary School Version

Please read the sentence and check the number that best fits what you think.

Check the best answer: 1-Never 2-Sometimes 3-Often 4-Always		
Trustworthiness	1. Kids in my class tell the truth.	1 2 3 4
	2. Kids in my class keep their promises.	1 2 3 4
	3. We trust each other in my class.	1 2 3 4
Respect	4. Kids in my class pick on each other. (R)	1 2 3 4
	5. Kids in my class behave.	1 2 3 4
	6. We respect each other in my class.	1 2 3 4
Responsibility	7. Kids in my class do the best work they can.	1 2 3 4
	8. Kids in my class stay out of trouble.	1 2 3 4
	9. The kids in my class are responsible.	1 2 3 4
Fairness	10. Kids in my class play fair.	1 2 3 4
	11. Kids share with each other in my class.	1 2 3 4
	12. We treat each other the same in this class.	1 2 3 4
Caring	13. We care about each other in my class.	1 2 3 4
	14. In my class we help each other out.	1 2 3 4
	15. Kids in my class are nice to someone who is sad.	1 2 3 4
Citizenship	16. Kids in my class follow the class rules.	1 2 3 4
	17. Kids in my class get along with each other.	1 2 3 4
	18. We listen to each other's ideas in my class.	1 2 3 4

APPENDIX III

Classroom Community Environment Survey (CCES) Middle/High School Version

Instructions

Please read the numbered items and select your best response. There are no right answers. What you think is what we're looking for.

Please select one response for each item below.					
	1-Almost Never	2-Rarely	3- Sometimes	4-Often	5-Almost Always
In my classes, students...					
1. Tell the truth.	1	2	3	4	5
2. Believe honesty is the best choice.	1	2	3	4	5
3. Are trustworthy.	1	2	3	4	5
4. Avoid cheating on school work.	1	2	3	4	5
5. Keep promises they make to other students.	1	2	3	4	5
6. Are loyal to other students.	1	2	3	4	5
7. Avoid lying to avoid consequences.	1	2	3	4	5
8. Are honest.	1	2	3	4	5
9. Are generally truthful.	1	2	3	4	5
In my classes, students...					
10. Treat each other with respect.	1	2	3	4	5
11. Respect the personal property of others.	1	2	3	4	5
12. Show respect for school property.	1	2	3	4	5
13. Behave respectfully toward school staff.	1	2	3	4	5
14. Show respect toward teachers.	1	2	3	4	5
15. Show good sportsmanship.	1	2	3	4	5
16. Are respectful toward each other.	1	2	3	4	5
17. Avoid put-downs.	1	2	3	4	5
18. Do not bully one another.	1	2	3	4	5
In my classes, students...					
19. Do not break promises.	1	2	3	4	5
20. Are not late to class.	1	2	3	4	5
21. Feel responsible in helping each other.	1	2	3	4	5
22. Get their school work done on time.	1	2	3	4	5
23. Take responsibility for their actions.	1	2	3	4	5
24. Do the best job they can in their school work.	1	2	3	4	5
25. Have self-control.	1	2	3	4	5
26. Report things that are wrong.	1	2	3	4	5
27. Generally know right from wrong.	1	2	3	4	5
In my classes, students...					
28. Are concerned with fairness.	1	2	3	4	5
29. Tend to be treat each other the same.	1	2	3	4	5
30. Are tolerant of one another.	1	2	3	4	5
31. Are equal to one another.	1	2	3	4	5
32. Are fair toward one another.	1	2	3	4	5
33. Consider fairness in working with each other.	1	2	3	4	5
34. Pay attention to things being equal for one another.	1	2	3	4	5
35. Are equal in their classroom interactions with each other.	1	2	3	4	5
36. Are open to other students' ideas.	1	2	3	4	5

In my classes, students...					
37. Make an effort to include those who are different.	1	2	3	4	5
38. Try to comfort peers who have experienced sadness.	1	2	3	4	5
39. Help each other, even if they are not friends.	1	2	3	4	5
40. Work well together.	1	2	3	4	5
41. Help new students feel accepted.	1	2	3	4	5
42. Are willing to forgive each other.	1	2	3	4	5
43. Are patient with each other.	1	2	3	4	5
44. Listen to each other in class discussions.	1	2	3	4	5
45. Share what they have with others.	1	2	3	4	5
In my classes, students...					
46. Do the right thing even if it's not in their best interest.	1	2	3	4	5
47. Resolve conflicts without fighting.	1	2	3	4	5
48. Try to get their friends to follow the rules.	1	2	3	4	5
49. Take an active role in solving classroom problems.	1	2	3	4	5
50. Consider each other's opinions.	1	2	3	4	5
51. Work together to bring about changes in the classroom community.	1	2	3	4	5
52. Obeys classroom rules.	1	2	3	4	5
53. Tend to give more than they take.	1	2	3	4	5
54. Have a say in classroom control.	1	2	3	4	5

APPENDIX IV

Service Learning Survey

Instructions

Then, answer the questions the best you can. There are no right or wrong answers, we just need your opinions and information.

1. Are you enrolled in the V.A.L.U. service learning class? Yes ___ No ___

2. How many hours a month do you do volunteer work in your community? _____

3. How often have you done volunteer work during the past year?

Not at all ___ Occasionally ___ Frequently ___

4. Since entering middle school, have you performed any community/volunteer service?

Yes ___ No ___

If yes, how was the service performed? (check as many as needed)

___ a. as part of a course or class

___ b. as part of a school-sponsored activity (example: National Honor Society, school club, school team, etc.).

___ c. as a part of a non-school group (example: church, family, etc.).

5. I consider myself a(n)... (check one)

___ A student

___ B student

___ C student

___ F student

Please circle one response for each item below.

1-Never 2-Seldom 3-Often 4-Always

By volunteering in the community I can learn...				
6. More about it.	1	2	3	4
7. About different careers.	1	2	3	4
8. More about myself.	1	2	3	4
9. About people who are different than me.	1	2	3	4
10. About other cultures.	1	2	3	4
11. More about responsibilities.	1	2	3	4
12. More about the working world.	1	2	3	4
13. New skills.	1	2	3	4
So far, during this school year...				
14. I had adult responsibilities	1	2	3	4
15. I made important decisions	1	2	3	4
16. I did things myself instead of observing	1	2	3	4
17. I had freedom to develop and use my own ideas	1	2	3	4
18. I had freedom to explore my own interests	1	2	3	4
19. My ideas were considered	1	2	3	4

So far, during this school year...				
20. I had challenging tasks	1	2	3	4
21. I felt I made a contribution	1	2	3	4
22. What I did was interesting	1	2	3	4
In Floresville...				
23. I feel like I am part of a community.	1	2	3	4
24. I feel like I can make a difference in the community.	1	2	3	4
25. I have a strong attachment to my community.	1	2	3	4
Generally...				
26. I feel that doing something to help others is important to me.	1	2	3	4
27. I like to help other people, even if it is hard work.	1	2	3	4
28. I know what I can do to help make the community a better place.	1	2	3	4
29 I feel I have the power to make a difference in the community.	1	2	3	4
30. I believe that I can personally make a difference in the community.	1	2	3	4
I believe...				
31. Helping other people is something everyone should do, including myself.	1	2	3	4
32. It is my responsibility to help improve the community.	1	2	3	4
33. I have a personal responsibility to contribute in some way to the community.	1	2	3	4
34. Helping other people is something I am personally responsible for.	1	2	3	4
35. Being actively involved in community issues is everyone's responsibility, including mine.	1	2	3	4

I think being active in my community is: <i>(circle the number closest to what you think)</i>					
36. Dumb	1	2	3	4	5 Smart
37. Unusual	1	2	3	4	5 Usual
38. Worthless	1	2	3	4	5 Valuable
39. Bad	1	2	3	4	5 Good
40. Not important	1	2	3	4	5 Important
41. Boring	1	2	3	4	5 Interesting
42. Old-fashioned	1	2	3	4	5 Modern
43. Selfish	1	2	3	4	5 Unselfish
44. Useless	1	2	3	4	5 Useful
45. Dishonest	1	2	3	4	5 Honest
46. Something I don't like	1	2	3	4	5 Something I like
I think doing volunteer/service work is: <i>(circle the number closest to what you think)</i>					
47. Dumb	1	2	3	4	5 Smart
48. Worthless	1	2	3	4	5 Valuable
49. Bad	1	2	3	4	5 Good
50. Not important	1	2	3	4	5 Important
51. Boring	1	2	3	4	5 Interesting
52. Old-fashioned	1	2	3	4	5 Modern
53. Selfish	1	2	3	4	5 Unselfish
54. Useless	1	2	3	4	5 Useful
55. Dishonest	1	2	3	4	5 Honest
56. Something I don't like	1	2	3	4	5 Something I like

APPENDIX V

Building Developmental Assets in School Communities

Please complete the following “test” to the best of your ability. It’s okay if you do not have all of the answers, you may skip to the next item if needed. We do not expect you to know all of this character education content. Just do your best.

1. **External assets** are made up of:
 - a. Trustworthiness; Loyalty; Helpfulness; Friendliness; Courteousness; Kindness
 - b. Support; Empowerment; Boundaries & Expectations; Constructive Use of Time
 - c. Commitment to Learning; Positive Values; Social Competencies; Positive Identity
 - d. Shared purpose & identity; Alignment of practices with desired outcomes; Having a voice/taking a stand

2. **Internal assets** are made up of:
 - a. Trustworthiness; Loyalty; Helpfulness; Friendliness; Courteousness; Kindness
 - b. Support; Empowerment; Boundaries & Expectations; Constructive Use of Time
 - c. Commitment to Learning; Positive Values; Social Competencies; Positive Identity
 - d. Shared purpose & identity; Alignment of practices with desired outcomes; Having a voice/taking a stand

3. **Support** can be practiced by which of the following?
 - a. Modeling courtesy
 - b. Asking students for advice
 - c. Telling students what to expect.
 - d. Coaching students in making decisions

4. **Social Competencies** can be practiced by which of the following?
 - a. Modeling courtesy
 - b. Asking students for advice
 - c. Telling students what to expect.
 - d. Coaching students in making decisions

5. **Empowerment** can be practiced by which of the following?
 - a. Modeling courtesy
 - b. Asking students for advice
 - c. Telling students what to expect.
 - d. Coaching students in making decisions

6. **Boundaries and Expectations** can be practiced by which of the following?
 - a. Modeling courtesy
 - b. Asking students for advice
 - c. Telling students what to expect.
 - d. Coaching students in making decisions

7. The primary developmental asset schools can most directly impact is/are:
- a. creative activities
 - b. conflict resolution
 - c. community values
 - d. achievement motivation
8. Which one of the following is a “new attitude” according to the Search Institute?
- a. Focus on student’s problems
 - b. Building developmental assets is a program
 - c. We’re already building assts
 - d. Young people are resources
9. How competent do you feel as a Developmental Assets teacher? *Please select one*
- | | | | | |
|--------------------|---|--------------------|---|----------------|
| 1 | 2 | 3 | 4 | 5 |
| Not very competent | | Somewhat competent | | Very competent |
10. Overall, how effective do you feel you are at teaching “character”? *Please select one*
- | | | | | |
|--------------------|---|--------------------|---|----------------|
| 1 | 2 | 3 | 4 | 5 |
| Not very effective | | Somewhat effective | | Very effective |

APPENDIX VI

Powerful Teaching

1. What is an asset-building lens?
 - a. A way of viewing student work
 - b. A visible curricular teaching aid
 - c. A way of viewing curriculum content
 - d. A visible teaching trait
2. List as many strategies as you can for infusing Developmental Assets in your classroom/content area:
3. What is the connection between building assets and improving student achievement?
4. How competent do you feel as an asset-building teacher? *Please select one*

1 2 3 4 5

Not very competent Somewhat competent Very competent

5. There are 8 pro-social asset categories with 4 of them representing external structures, relationships, and activities that create a positive environment for young people. The other 4 categories reflect internal values, skills, and beliefs young people need to fully engage and function.

Please read over the following list and select whether the category focuses on external or internal assets.

Asset Category	External Assets	Internal Assets
Boundaries & expectations	___	___
Commitment to learning	___	___
Constructive use of time	___	___
Empowerment	___	___
Positive identity	___	___
Positive values	___	___
Social competencies	___	___
Support	___	___

6. We know that asset building in classrooms and schools can occur in three key areas. For each area, offer a tangible example of what you are doing/could do to focus more intentionally on asset building:
 - a. Relationships:
 - b. The classroom environment:
 - c. Your classroom curriculum:

d. Your instructional practice:

7. Teachers have the power to positively influence student behavior by engaging with them at 6 levels. For each level, offer a tangible example of what you do/could do to engage with a student.

a. Level 1 (Making Connections with Young People):

b. Level 2 (Getting to Know Them):

c. Level 3 (Identifying and Encouraging Strengths and Talents):

d. Level 4 (Challenging To Be and Do Their Best):

e. Level 5 (Focusing on Building Specific Assets):

f. Level 6 (Maintaining the Relationship):

APPENDIX VII

Smart and Good High Schools

1. Identify the six principles by which an ethical learning community is created.
 - a. (1) Trustworthy; (2) Loyal; (3) Helpful; (4) Friendly; (5) Courteous; (6) Kind.
 - b. (1) Shared purpose & identity; (2) Align practices with desired outcomes; (3) Have a voice/take a stand; (4) Take personal responsibility; (5) Practice collective responsibility; (6) Deal with the tough issues.
 - c. (1) Trustworthiness; (2) Respect; (3) Responsible; (4) Fair; (5) Caring; (6) Friendly.
 - d. (1) Be proactive; (2) Begin with the end in mind; (3) Put first things first; (4) Think win/win; (5) Seek first to understand, then to be understood; (6) Synergize.
2. *Performance* and *performance character* are the same thing. T F
3. *Task orientation* is when a student seeks to:
 - a. complete their homework
 - b. surpass their own performance
 - c. surpass other's performance
 - d. complete tasks to please the teacher
4. *Ego orientation* is when a student seeks to:
 - a. complete their homework
 - b. surpass their own performance
 - c. surpass other's performance
 - d. complete tasks to please the teacher
5. The notion of *moral character* is when students:
 - a. help others out at school
 - b. compete to be the best at being "good"
 - c. diminish their task orientation in favor of ego orientation
 - d. use ethical means to achieve performance goals
6. A Professional Ethical Learning Community (PELC) consists of:
 - a. all school staff
 - b. local doctors, lawyers, accountants, and school principals
 - c. parents, teachers, students
 - d. community leaders and school administrators
7. List specific things you do to promote an ethical learning community in your classroom.

8. Which of the following are considered promising practices for developing diligent and capable performing students?

- a. foster peer relations; teach the art of asking questions.
- b. use rubrics to help students self-assess; develop high-challenge rites of passage.
- c. model ethical thinking; help students develop a 'go-to' framework.
- d. develop a student voice in the school; promote on-going self-reflection in students.

9. Which of the following are considered promising practices for developing an ethical learning community?

- a. foster peer relations; teach the art of asking questions.
- b. use rubrics to help students self-assess; develop high-challenge rites of passage.
- c. model ethical thinking; help students develop a 'go-to' framework.
- d. develop a student voice in the school; promote on-going self-reflection in students.

10. How competent do you feel as an ethical learning community teacher? *Please select one*

1	2	3	4	5
Not very competent		Somewhat competent		Very competent

11. Overall, how effective do you feel you are at teaching character? *Please select one*

1	2	3	4	5
Not very effective		Somewhat effective		Very effective

APPENDIX VIII

High School Transition Survey (HSTS)

Instructions

Insert your student ID number in the box below. Then, circle the number form 1-4 that best describes how you feel. There are no right or wrong answers, we simply want your opinion.

Please select one response for each item below. 1-Never 2-Seldom 3-Often 4-Always	
1. I feel close to students at my school.	1 2 3 4
2. I feel close to teachers at my school.	1 2 3 4
3. I feel like I am part of my school.	1 2 3 4
4. I am happy to be at my school.	1 2 3 4
5. Teachers treat students fair.	1 2 3 4
6. I feel safe in my school.	1 2 3 4
7. I get along with my teachers.	1 2 3 4
8. I get along with other students.	1 2 3 4
9. I feel like my teachers care about me.	1 2 3 4
10. When I am in class, I work hard.	1 2 3 4
11. I can get good grades in school.	1 2 3 4
12. I am smart when it comes to schoolwork.	1 2 3 4
13. I can do well in school if I want to.	1 2 3 4
14. I try hard in my schoolwork.	1 2 3 4
15. I get good grades no matter what I do.	1 2 3 4
16. I think I am pretty smart in school.	1 2 3 4
17. If I decide to learn something difficult, I can.	1 2 3 4
18. I can accomplish almost anything if I try.	1 2 3 4

APPENDIX IX

Year 1 LACE Test

Instructions: Please select the best answer for each item below, then press the Submit button.

1. Which of the following is the most significant ingredient in successful school reform according to Marvin Berkowitz?

- a. teacher attitudes
- b. leadership
- c. length of the school day
- d. state funding

2. Character education success depends upon it becoming which level of priority with school principles?

- a. second priority after students
- b. just behind TAKS scores
- c. the highest priority
- d. whatever priority David Vinson says

3. According to Turning Points 2000, the role of an effective principal is to...

- a. act as a change agent
- b. create a culture of character
- c. manage his/her school well
- d. be certain TAKS scores continually improve

4. Which of the following lists how principals can foster teacher leadership?

- a. Act, Create, Manage, Improve
- b. Hold accountable, Request, Provide, Make the tough decisions, Collaborate, Trust
- c. Anticipate, Let go, Trust, Authorize, Involve, Shield, Impart, Provide
- d. Expect, Relinquish, Entrust, Empower, Include, Protect, Share responsibility, Give recognition

5. Berkowitz' "Moral Anatomy" consists of which of the following lists?

- a. Trustworthy, loyal, helpful, friendly, courteous, kind
- b. Action, values, personality, emotion, reasoning, identity, and foundational characteristics
- c. Trustworthiness, respect, moral responsibility, fairness, caring, and friendliness
- d. Be proactive, begin with the end in mind, put first things first, think win/win, seek first to understand, then to be understood, synergize.

6. Read the following passage, then select the best answer.

"I've come to the frightening conclusion that I am the decisive element in the classroom. My personal approach creates the climate. My daily mood makes the weather. As a teacher, I possess a tremendous power to make a child's life miserable or joyous. I can be a tool of torture or an instrument of inspiration. I can humiliate or humor, hurt or heal. In all situations, it is my response that decides whether a crisis will be escalated or de-escalated and a child humanized or dehumanized." –Dr. Haim Ginott (1922-1973)

This passage supports Berkowitz' statement on...

- a. the 1970s perspective of educational psychology
- b. why teachers should foster character development
- c. the psychological power of principals

d. communication approaches school leaders must be aware of

7. The number one influence on a child's character is?

- a. religious affiliation
- b. peers
- c. unrelated significant adults
- d. parents

8. From a historical perspective, character education...

- a. results from No Child Left Behind
- b. largely vanished after World War II
- c. was most prominent just after World War II
- d. has always been agreed upon by educators

9. In recent research, character education has been demonstrated to...

- a. take significant time away from academic study in middle schools
- b. be too controversial to be implemented well in public education
- c. marginally improve student achievement in science
- d. potentially improve student academic achievement

10. Berkowitz notes that character education is which of the following?

- a. "Rocket science" and school reform
- b. Relatively easy to implement if schools follow the 11 character education principles (CEPs)
- c. A reactionary result of today's society
- d. The best method for improving student scores on standardized tests

11. Research demonstrates that the most prominent outcome of character education programs is?

- a. reduced drug use by students
- b. increased student attitudes toward school
- c. improved psycho-social learning environments
- d. improved student socio-moral cognition

12. How *competent* do you feel as a character education leader? *Please select one*

1	2	3	4	5
Not very competent		Somewhat competent	Very competent	

13. Overall, how *effective* do you feel you are as character education leader? *Please select one*

1	2	3	4	5
Not very effective		Somewhat effective	Very effective	