Research and Accountability Department

EVALUATION REPORT

BUREAU OF PROGRAM EVALUATION

Volume 7, Issue 2, September 2013

The impact of the Baylor College of Medicine (BCM) 2011 and 2012 Summer Institute on elementary and middle-school students' science performance

By Venita Holmes, Dr.P.H.

The purpose of this evaluation was to measure the impact of the BCM Summer Institute on science performance of students whose teachers participated during the summers of 2011 and 2012. Paired t-tests revealed statistically significant increases in a sample of elementary students' Stanford 10 science scores at 3rd through 5th grades in 2012 and at 5th grade in 2013. For the middle-school student sample, significant increases were found at 7th and 8th grades in 2012 and 2013. Linear mixed-effects modeling revealed elementary students' previous year Stanford 10 science scores and teachers' additional hours of professional development were the strongest predictors of students' post BCM teacher participation science performance in both years. Propensity score matching yielded increases in elementary and secondary students' STAAR scores following teacher participation relative to comparison group students; however, the differences were not statistically significant at p <. 05. The level of mentoring and peer support that teachers received during the school year was not measured in this evaluation and could have accounted for differences in science performance outcomes at elementary and secondary levels.

Background

The Baylor College of Medicine (BCM) Summer Science Institute presents current and effective teaching strategies, in-depth science content lessons, assessments, and relevant reading and mathematics concepts that are aligned with TEKS and STAAR recommended objectives (BCM, 2013). This professional development (PD) program targeted elementary (prekindergarten - 5th grades) and secondary (6th - 8th grades) HISD teachers who taught science in self-contained classes, science labs, and regular education classes. The program recruited teachers who were new to science instruction or needed additional support in their teaching. The Institute helped teachers develop science knowledge, teaching skills, and confidence needed to prepare their students in science and math (BCM, 2013). Elementary teachers were offered 15 days of professional development (10 days during the summer; five half-days during the school year). Secondary teachers were engaged in 8 days of professional development (four in the summer; four half-days during the school year). Gifted/talented and continuing education credits along with stipends were offered for participation in the Institute.

Research on teacher education maintains that effective science instructors must know how to teach science and how students learn science (Cotabish, Dailey, Hughes, and Robinson, 2011). Becoming an effective science teacher is a continuous process that extends beyond undergraduate education throughout a teacher's professional career. Science continuously changes relative to knowledge and relevance to society. Teachers need ongoing professional development opportunities to build their science knowledge and skills and their understanding of students' diverse interests, abilities, and experiences to successfully support and guide student learning (Nadelson, et al., 2012; National Academies of Science, 2013). Professional development provides an opportunity for teachers to study and engage in research on science teaching and learning, and to share with colleagues what they have learned (National Academies of Science, 2013, p. 55).

Loucks-Horsley et al. (2003) proposed a theoretical model for designing science professional development. The model relies on a decision-making process that is informed by relevant inputs during the planning process, a commitment to vision and standards, analysis of student learning, goal setting, planning, doing, and evaluating program outcomes. It is important that teachers have a knowledge base and beliefs of the unique aspects of the model and incorporate PD strategies in the delivery of content.

Several studies have been conducted on science professional development incorporating a summer institute model. Nadelson et al. (2012) conducted a study of a program designed to enhance teacher capacity and effectiveness to teach science. Teachers focused on science content knowledge and affective perceptions in the context of teaching and learning. The study found significant gains in the participating teachers' perceived efficacy, comfort, contentment, and knowledge related to science (Nadelson, et al. p. 81). A randomized controlled trial measuring teacher and student learning in science (Cotabish et al., 2010; Dailey et al., 2012) found that after a 2-year intervention consisting of 120 hours of professional development during a summer institute and one-to-one peer coaching, there was a statistically significant gain in science process skills by elementary teachers in the experimental group compared with teachers in the control group. Experimental teachers demonstrated increased confidence in their ability to lead students in developing science process skills. Sinclair, Naizer, & Ledbetter (2011) reported that teachers with increased confidence in science teaching were more likely to encourage inquiry-learning in their classroom. Furthermore, Liu, Lee, and Lim (2010) reported that positive teacher perceptions about inquiry-based instruction were associated with improved student achievement. Teacher beliefs and attitudes in science teaching, particularly the use of inquiry-based learning, had a considerable impact on classroom instruction and student learning (Choi & Ramsey, 2009).

There is limited research on the extent that teacher professional development in science influences students' science achievement. This study will add to the body of knowledge in this area.

Methodology

Based on databases provided by BCM administrative staff, 266 elementary and 147 secondary science teachers were identified as participants in the 2011 BCM Summer Institute. The 2012 Institute reportedly consisted of 265 elementary and 69 secondary teachers. The number of hours teachers attended the summer component of the program and additional professional development hours during the school year were also provided BCM program staff. Data on only teachers who participated in the summer component of the program were included in the analysis (**Appendix A**).

Data Analysis

BCM Summer Institute teacher databases were linked to the Chancery database to obtain student samples. Mean hours of professional development were calculated for students who had multiple science teachers. Professional development hours by Stanford 10 environment/science results for students of 2011 and 2012 teachers were extracted from the test databases. Paired t-test analysis was conducted for students with two years of Stanford 10 science data as a pre/post-teacher participation measure. STAAR results at tested grades 5 and 8 were also analyzed in this evaluation.

Linear Mixed-Effects Modeling

Linear mixed-effects modeling was used to analyze how much the variation from teacher to teacher accounted for student's 2012 and 2013 Stanford 10 performance following teacher participation in the program. Preliminary analyses were performed to ensure no violation of the assumptions of normality, linearity, multicollinearity, and homoscedasticity. The model controlled for summer professional development hours, additional professional development during the regular school year, and students' previous year Stanford 10 science scores. Gender and economic status were initially tested as control variables. These variables did not make statistically significant contributions and were deleted from the models. Only students with previous year Stanford 10 science scores were included in the analysis.

Propensity Score Matching

Propensity score matching (PSM) and "nearest-neighbor matching" techniques were also used to assess the impact of the BCM Summer Science Institute on science achievement to develop treatment and comparison groups. PSM is considered a viable method to estimate causal treatment effects, selection bias, and control for observed bias (Rosenbaum & Rubin, 1985; Rubin, 1997; Joffe & Rosenbaum, 1999; Murname and Willett, 2011). Stata software was used to select comparison groups. When more than one good match existed for the treatment group, Stata's "attnd algorithm" randomly picked one of the duplicates with replacement to become the actual

Table 2: Elementary Students' 2012 Stanford 10 Performance following Teacher Participation in 2011 BCM Summer Institute

	n	Mean NCE	Std. Devia.	†District Mean
Enviro.				
1st	1193	49	22.2	49
2nd	1639	53	23.3	53
Science				
3rd	1513	56	22.1	53
4th	1802	52	20.4	51
5th	2555	61	23.2	61

Note: 34.4-64.9 NCEs within the average range (Stanford Achievement Test, 1997); †All students

neighbors (Murname and Willett, 2011). Stata identified some students as matches for multiple treatment group students.

Propensity score matching criteria consisted of students' gender, at risk status, previous science performance on Stanford 10, and gifted/talented status. Pre-science performance was centered on the group means. The variables were used as fixed effects and random effects to adjust for the variation within and between groups. These variables were selected because they were considered social factors that could have an effect on student performance (Beckar and Luthar, 2002).

What was the impact of the 2011 BCM Summer Institute on student science performance?

Stanford 10 and STAAR science test results were used to measure the impact of the BCM Summer Institute on student science performance. The results are presented by elementary and secondary levels for students of teachers who attended the BCM 2011 Summer Science Institute.

Summer 2011 Elementary Level Teacher/Student Sample

The findings in **Table 2** reflect the spring 2012 Stanford 10 results of students following their teachers' participation in the 2011 BCM Summer Institute. Stanford 10 test results for 2,832 students in grades 1 and 2 on the environment subtest and 5,870 students in grades 3 through 5 on the science subtest are presented. District means are included for comparison.

Table 2 shows that at the elementary level, 1st through 5th-grade student performance fell within the average range on the Stanford 10 environment and science subtests in 2012 (between 34.4-64.9 NCEs). Second-grade students attained a higher mean NCE than 1st-grade students on the environment subtest (52.8 NCEs vs. 49.3 NCEs). Fifth grade student performance fell within the average on the science subtest (61.0 NCEs).

Table 3 presents the results of linear mixed effects modeling based on 2012 Stanford 10 results as the predictor variable and summer professional development hours, additional professional development hours during the school year, and 2011 Stanford 10 science subtest NCEs as covariates in the model. *Beta* coefficients and *p*-values are depicted. The findings revealed that students' Stanford 10 scores in 2011 and additional professional development during the school year were strong predictors of 2012 Stanford 10 science performance for the elementary student sample ($p < .001$ and $p < .037$, respectively).

Table 3. Elementary Student Linear Mixed-Effects Modeling Predicting 2012 Stanford Science

	β	*p*
2011 Stanford Science NCEs	.744	.000**
Summer PD Hours	4.73	.522
Additional PD Hours	7.24	.037*

Sample represents 14 teachers, 26 schools, 1058 students of 2011 Summer PD.
Note: *p < .05, ** p < .001
2011 Mean NCE = 50.8; 2012 NCE Mean NCE= 57.0

Table 4 shows the results of a paired t-test to assess the impact of the program on Stanford 10 science scores for students with two years of test results (spring 2011 and spring 2012). There were 1,058 students who met this criterion. Twenty-six elementary schools were represented in the data and 13 elementary school teachers. Some teachers were linked to more than one school. Based on the student sample, there was a statistically significant increase in science scores from pre- to posttest at

Table 4. Elementary Student Sample Paired T-test Analysis for Students with Two Years of Stanford 10 Science Results, 2011 and 2012

Grade Level		Test Year	n	Mean	Std. Devia.	Mean Diff.	t	p (one-tailed)
3rd	Pretest	2011	12	23.7	10.3	14.5	4.15	.001*
	Posttest	2012	12	38.2	8.7			
4th	Pretest	2011	301	48.6	19.5	1.5	1.94	.027*
	Posttest	2012	301	50.1	19.0			
5th	Pretest	2011	745	52.1	19.8	8.0	12.15	.000*
	Posttest	2012	745	60.1	22.8			

Table 5: Secondary Students' 2012 Stanford 10 Performance following Teacher Participation in 2011 BCM Summer Institute

	n	Mean NCE	Std. Devia.	†District Mean
Grade Level				
6th	2066	50	19.0	49
7th	4094	56	19.3	56
8th	4042	54	18.7	56

Note: 34.4-64.9 NCEs are within the average range (Stanford Achievement Test, 1997); †All students

grades 3,4, and 5. The highest mean difference in scores was noted at 3rd grade by 14.5 NCEs.

Summer 2011 Secondary Teacher/Student Sample

Stanford 10 science subtest results of 10,202 secondary students in grades 6 through 8 whose teachers participated in the 2011 BCM Summer Institute are presented in **Table 5**. The mean NCE at each secondary grade level fell within the average range (between 34.4 - 64.9 NCEs). The 7th-grade student sample had the highest performance (56 NCEs) compared to the 6th and 8th-grade student samples (50 and 54 NCEs, respectively).

Table 6 shows the *Beta* coefficients and *p*-values derived from linear mixed-effects modeling based on 2012 Stanford 10 results as the predictor variable and summer professional development hours and 2011 Stanford 10 science NCEs as covariates in the model. The model shows that students' 2011 Stanford 10 science performance was a strong predictor of 2012 science performance on the Stanford 10 science subtest ($p < .05$). Summer professional development hours of teachers did not contribute significantly to the model.

A paired t-test analysis was conducted based on Stanford 10 science results of the 2012 secondary student group with two years of test data. The spring 2011 science scores was the pretest measure and the 2012 science test scores was the posttest measure. A total of 8,969 students were included in the analysis, representing 75 secondary teachers at 41 schools. The results are presented in **Table 7**. There was a statistically significant increase in the performance of students in grades 7 and 8 and a statistically significant decrease in students' performance at grade 6 ($p < .001$). The largest mean difference over the two-year period was at 6th grade (-8.3 NCEs).

Table 6. Secondary Student Linear Mixed-Effects Modeling Predicting 2012 Stanford Science

	β	p
2011Stanford Science NCEs	.352	.007*
Summer PD Hours	.134	.612

Sample represents 75 teachers, 49 schools, 8969 students of 2011 Summer PD.
Note: one-tailed test; *p < .05
2011 Mean NCE = 53.0; 2012 NCE Mean NCE= 54.5

How did students whose teachers participated in the 2012 BCM Summer Institute perform in science?

Summer 2012 Elementary Student/Teacher Sample

The outcomes in **Table 8** represent the spring 2013 Stanford 10 environment and science subtests results of 8,020 elementary students. These scores were gathered in the spring following teacher participation in the BCM Summer Institute. The 2nd-grade student sample had a higher mean score than the 1st-grade student sample on the environment subtest (46.9 vs. 42.3). Fifth-grade students had a higher mean score than 3rd and 4th-grade students on the science subtest (54.1 NCEs compared to 47.4 NCEs and 50.8 NCEs).

Linear mixed effects modeling was conducted to determine the impact of the program on elementary students' 2013 science performance. The results are presented in **Table 9** for 2003 students of 38 teachers at 29 schools. Students' 2013 Stanford 10 NCEs were modeled as the predictor variable while

Table 7. Secondary Student Sample Paired T-test Analysis for Students with Two Years of Stanford 10 Science Results, 2011 and 2012

Grade Level		Test Year	n	Mean	Std. Devia.	Mean Diff.	t	p (one-tailed)
6th	Pretest	2011	1792	58.5	20.2	-8.3	-22.96	.000*
	Posttest	2012	1792	50.2	18.8			
7th	Pretest	2011	3602	52.6	21.1	3.9	15.56	.000*
	Posttest	2012	3602	56.5	19.0			
8th	Pretest	2011	3575	50.7	20.7	4.0	16.77	.000*
	Posttest	2012	3575	54.7	18.3			

Table 8: Elementary Students' 2013 Stanford 10 Performance following Teacher Participation in 2012 BCM Summer Institute

	n	Mean NCE	Std. Devia.	†District Mean
Environ.				
1st	845	42	22.0	47
2nd	1002	47	20.5	50
Science				
3rd	1316	47	20.8	51
4th	1942	51	21.0	52
5th	2915	54	20.6	55

Note: 34.4-64.9 NCEs within the average (Stanford Achievement Test, 1997); †All students

Table 9. Elementary Student Sample Linear Mixed-Effects Modeling Predicting 2013 Stanford Science

Grades 3 - 5	β	p
2012 Stanford Science NCEs	1.43	.000**
Summer PD Hours	.79	.037*
Additional PD Hours	8.91	.001*

Sample represents 38 teachers, 29 schools, 2003 students of 2012 Summer PD.
Note: one-tailed test; *p < .05, ** p < .001
2012 Mean NCE = 52.7; 2013 NCE Mean NCE= 55.01

Table 10. Elementary Student Sample Paired T-test Analysis for Students with Two Years of Stanford 10 Science Results, 2012 and 2013

Grade Level		Test Year	n	Mean	Std. Devia.	Mean Diff.	t	p (one-tailed)
3rd	Pretest	2012	15	36.2	12.3	-4.3	-.244	.406
	Posttest	2013	15	31.9	12.6			
4th	Pretest	2011	695	54.8	19.8	-0.1	-.114	.455
	Posttest	2012	695	54.7	21.7			
5th	Pretest	2011	1293	51.7	19.2	3.7	8.94	.000*
	Posttest	2012	1293	55.4	20.8			

summer professional development and additional hours of professional development during the school year were covariates in the model. *Beta* coefficients and *p*-values are depicted. The findings in Table 9 show that previous year Stanford 10 performance ($p < .001$), summer professional development ($p < .05$), and additional professional development ($p< .001$) were strong predictors of 2013 Stanford performance.

A paired t-test analysis was conducted that included the Stanford 10 science results of students with two years of data. The 2013 test results were analyzed as the posttest and the 2012 test results were the pretest measure. **Table 10** shows that mean scores of 5th grade students increased from pretest (2012) to posttest (2013). There was a statistically significant increase in the scores of the 5th-grade student sample from 2012 to 2013 ($p < .001$). The decreases in grades 3 and 4 were not statistically significant.

Summer 2012 Secondary Student/Teacher Sample

The 2013 Stanford 10 science subtest results of secondary students in grades 6 through 8 whose teachers participated in the 2012 BCM Summer Institute are presented in **Table 11**. Eight-grade students had the highest mean science NCE (53.0), followed by 6th and 7th-grade students (46.4 and 45.1 NCEs).

Multilevel modeling resulted in data for 2,866 students of 31 secondary teachers at 18 schools. The results are presented in **Table 12**. Students' previous year 2012 Stanford 10 science scores

Table 11: Secondary Students' 2013 Stanford 10 Performance following Teacher Participation in 2012 BCM Summer Institute

	n	Mean NCE	Std. Dev	†District Mean
Grade				
6th	3036	46.4	20.4	52
7th	2573	45.1	20.3	51
8th	2371	53.0	18.1	57

Note: 34.4 - 64.9 NCEs within the average range (Stanford Achievement Test, 1997); †All students

Table 12. Secondary Student Sample Linear Mixed-Effects Modeling Predicting 2013 Stanford Science

Grades 6 – 8	β	p
2012 Stanford Science NCEs	1.17	.102
Summer PD Hours	3.12	.102
Additional PD Hours	6.29	.133

Sample represents 31 teachers, 18 schools, 2866 students of 2012 Summer PD.
Note: one-tailed test; *p < .05, ** p < .001
2012 Mean NCE = 50.5; 2013 NCE Mean NCE= 51.9**

Table 13. Secondary Student Sample Paired T-test Analysis for Students with Two Years of Stanford 10 Science Results, 2012 and 2013

Grade Level		Test Year	n	Mean	Std. Devia.	Mean Diff.	t	p (one-tailed)
6th	Pretest	2012	1170	56.4	20.4	-9.7	-16.39	.000*
	Posttest	2013	1170	46.7	20.1			
7th	Pretest	2012	689	44.3	17.2	1.8	3.01	.003*
	Posttest	2013	689	46.1	19.6			
8th	Pretest	2012	1007	52.0	17.5	3.6	2.79	.000*
	Posttest	2013	1007	55.6	17.6			

were the strongest predictors of their Stanford 10 science performance in 2013.

A paired t-test analysis was conducted consisting of the Stanford 10 results of a sample of secondary students with two years of test data (**Table 13**). Students' performance in spring 2013 on the test represented the posttest score and the 2012 science performance was the pretest score. The results of 6,954 students were included in the analysis. There was a statistically significant increase in the performance of students in grades 7th by 1.8 NCEs and 8 by 3.6 NCEs ($p < .001$). However, the students' performance at 6th grade decreased by 9 NCEs. These results were also statistically significant ($p < .001$).

What was the effect of teacher participation in the BCM Summer Institute on students' STAAR science test performance?

STAAR science test results were analyzed for 5th and 8th-grade students whose teachers participated in the 2012 BCM Summer Institute. The STAAR scale scores and "advanced" level performance based on the phase-in 1 standard were utilized in the analysis.

Table 14 shows that the mean STAAR performance of the 5th and 8th–grade student samples fell within the "satisfactory" range. The mean score for 5th-grade students was slightly higher than the performance of the 8th grade student sample. In addition, the percent of 5th graders scoring at the "advanced level" on STAAR was slightly higher than the performance of 8th-grade students.

Comparison-Group Analysis for Summer 2012 Elementary Student/Teacher Sample

Propensity score matching yielded a comparison group relatively close to the treatment group (students whose teachers attended the 2012 Institute) on key fixed factors, including gender, economic status (free lunch), at risk status, and gifted/talented (g/t) classification. In addition, students were matched using 2012 Stanford 10 science results as a pre-science achievement measure although STAAR science results were the outcome measure.

Pearson correlation was conducted to determine whether there was a relationship between Stanford 10 science results and STAAR results for the 5th-grade student sample given that Stanford 10 results were used as the pre-science measure. Correlation analysis found a strong, positive correlation between 2012 Stanford 10 science subtest results and STAAR science performance of the 5th-grade student sample, $r = .695$, $n = 4533$, $p < .001$. **Appendix B** shows the results of matching on the fixed variables for 5th-grade students. Pre-science achievement matching results are presented in **Table 15**. (*Stata* software identified some matched students as controls for multiple treatment-group students.)

Table 15 shows that the treatment and comparison groups did not differ significantly on the Stanford 10 pre-science performance measure used in the propensity score matching. However, the treatment group had a slightly higher mean scale score on the STAAR science test following their teachers' participation in the BCM Summer Institute. The difference was not statistically significant at $p < .05$.

Table 14. STAAR Grade 5 and 8 Results, 2013

	Mean Scale Score	Std. devia.	% Advanced
5th Grade (n = 2818)	3660.10	62.7	6.7
**8th Grade (n = 2215)	3638.98	62.2	5.5

*5th grade – "Satisfactory" Scale Score": 3500- 4302
**8th grade – "Satisfactory" Scale Score: 3500-4355 (Texas Education Agency, 2013)

Table 15: **Grade 5** - 2013 STAAR Science Outcomes using Propensity Score Matching based on 2012 Stanford 10 Science NCEs, Gender, At Risk, and Gifted/Talented Status of BCM and Comparison Student Groups

	Pretest Science Performance Measure Stanford 10 Science 2012					**Posttest Science Performance Measure STAAR Science 2013**				
Grade 5	**Mean NCE**	**Std. Devia.**	**Mean Diff.**	**t**	**p**	**Mean Scale Score**	**Std. Devia.**	**Mean Diff.**	**t**	**p**
BCM	51.0	18.8	0.9	1.63	.103	3688.69	467.4	14.86	1.08	.277
Comparison	50.1	18.5				3673.83	455.5			

BCM group sample size = 2152; Comparison group sample size = 2381
Note: *Stata* software identified multiple students as comparisons of BCM students.

Table 16: **Grade 8** - 2013 STAAR Science Outcomes using Propensity Score Matching based on 2012 Stanford 10 Science NCEs, Gender, At Risk, and Gifted/Talented Status of BCM and Comparison Student Groups

	Stanford 10 Science Before BCM Science Institute 2012					**STAAR Science After BCM Science Institute 2013**				
Grade 8	**Mean NCE**	**Std. Devia.**	**Mean Diff.**	**t**	**p**	**Mean Scale Score**	**Std. Devia.**	**Mean Diff.**	**t**	**p**
Treatment	40.1	14.8	.1	.215	.830	3631.81	445.2	5.14	.239	.811
Control	40.0	14.8				3626.67	434.3			

BCM group sample size = 589; Comparison group sample size = 1369
Note: *Stata* software identified multiple students as comparisons of BCM students.

Comparison-Group Analysis for Summer 2012 Secondary Student/Teacher Sample

Stata software propensity score was used to form 8th-grade treatment and comparison student groups on key fixed factors. The groups are depicted in **Appendix C**. (*Stata* used some matched students as controls for multiple treatment-group students.) The 2012 Stanford 10 science subtest results of 8th-grade students were used as the pre-science performance measure. Pearson correlation was conducted to determine whether there was a relationship between Stanford 10 science results and STAAR results for the 8th-grade student sample. Correlation analysis found a strong, positive correlation between 2012 Stanford 10 science subtest results and STAAR science performance of the 8th-grade student sample, $r = .370$, $n = 1958$, $p < .001$.

Table 16 shows that treatment and comparison group students did not differ significantly in Stanford 10 science performance before or after teacher participation in the BCM Summer Institute.

Discussion

This study sought to assess the impact of the Baylor College of Medicine (BCM) Summer Institute on science performance of students whose teachers participated in the Institute. The scope of the evaluation encompasses elementary and secondary HISD teachers who attended the Institute during the summers of 2011 and 2012. BCM staff collaborated with HISD Professional Development staff and administrators in the Department of Curriculum, Instruction, and Assessment to align program content to state curriculum guidelines.

Summer professional development models have been used in educational settings in the past; however, little research has been done to explore impact on student academic performance.

To control for bias in measuring student science outcomes, matched comparison groups were formed, which included students at the same campuses as BCM Institute teachers. The study also used pre- and post-science measures to detect change in students' science achievement over time. The study found increases in performance on the norm-referenced Stanford 10 test from pre- to post teacher participation in the program for a sample of elementary and secondary students at most grade levels. Matched groups of elementary and secondary students revealed higher STAAR science scores for the BCM student group than the comparison group, however, differences were not significant.

There were several limitations to the study. First, STAAR and Stanford 10 science tests were used as the "best" measure of teacher impact on student performance, although these tests could be considered indirect measures of student outcomes. Second, differences in science performance for elementary compared to secondary student samples were observed; however, key moderating variables that could have influenced differences were not measured in the study. These moderating variables include the level of knowledge teachers acquired in science content during PD and the extent teachers applied knowledge in the classroom (e.g., instructional delivery).

Future research could include an examination of student performance on specific STAAR and Stanford 10 science test objectives that are aligned with the BCM Summer Institute PD content. This method could provide a more direct measure of program impact. Also, future research could investigate the influence of teacher knowledge and skills demonstrated on teacher assessments and through classroom observations on students' science performance.

References

Choi, S., & Ramsey, J. (2009). Constructing elementary teachers' beliefs, attitudes, and practical knowledge through an inquiry-based elementary science course. School Science and Mathematics, 109, 313-324. doi:10.1111/j.1949-8594.2009.tb18101.x

Cotabish, A., Dailey, D., Hughes, G. D., & Robinson, A. (2011). The effects of a STEM professional development intervention on elementary teachers' science process skills. Research in the Schools, 18(2), 16-25. Retrieved from http://search.proquest.com/docview/1284539880?accountid=35812

Cotabish, A., Robinson, A., MacFarlane, B., & Wood, B. (2010, April). Implementation of a gifted education STEM project: Teachers' initial perceptions, instructional behaviors, and knowledge of science content. Paper presented at the annual meeting of the American Education Research Association, Denver, CO.

Dailey, D., Cotabish, A., Robinson, A., & Hughes, G. (2012, April). Effects of implementing a STEM initiative on elementary teacher perceptions and concerns about science teaching and learning. Paper presented at the annual meeting of the American Education Research Association, Vancouver, BC.

Liu, O. L., Lee, H. S., & Linn, M. C. (2010). An Investigation of teacher impact on student inquiry science performance using a hierarchical linear model. Journal of Research in Science Teaching, 47, 807-819. doi:10.1002/tea.2037

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). Designing professional development for teachers of science and mathematics. (3rd ed.). Thousand Oaks, CA: Corwin Press. doi: 10.4135/9781452219103

Nadelson, L. S., Seifert, A., Moll, A. J., & Coats, B. (2012). i-STEM summer institute: An integrated approach to teacher professional development in STEM. Journal of STEM Education : Innovations and Research, 13(2), 69-83. Retrieved from http://search.proquest.com/docview/1009901964?accountid=35812

National Academies of Science (2013). National Science Education Standards: http://www.nationalacademies.org/rise/backg4.htm

National Research Council. (2007). Taking Science to School: Learning and Teaching Science in Grades K-8. Committee on Science Learning, Kindergarten Through Eighth Grade.

National Science Board. (2010). Preparing the next generation of STEM innovators: Identifying and developing our nation's human capital (NSB-10-33). Retrieved from http://www.nsf.gov/nsb/publications/2010/nsb1033

Sinclair, B. B., Naizer, G., & Ledbetter, C. (2011). Observed implementation of a science professional development program for K-8 classrooms. Journal of Science Teacher Education, 22, 579-594. doi:10.1007/s10972-010-9206-z

Stanford Achievement Test (1997). Technical Data Report. San Antonio, Texas: Harcourt Brace.

For additional information contact the HISD Department of Research and Accountability at 713-556-6700 or e-mail Research@Houstonisd.org.

Appendix A

Elementary 2011 Summer PD

Grade Level		N	Minimum	Maximum	Mean	Std. Deviation
03	**Summer PD**	12	57.25	65.00	63.7083	3.01668
	PD During School Year	12	5	20	14.17	7.017
04	**Summer PD**	301	57.25	65.00	63.5581	3.02089
	PD During School Year	301	5	20	15.98	5.653
05	**Summer PD**	745	57.25	65.00	63.8708	2.65507
	PD During School Year	745	5	20	15.26	5.709

Secondary 2011 Summer PD

Grade Level		N	Minimum	Maximum	Mean	Std. Deviation
06	**PD During School Year**	1170	.000	15.000	6.85011	5.716591
	Summer PD	1170	12.0	24.0	22.436	3.3273
07	**PD During School Year**	689	.000	20.000	11.50943	6.283941
	Summer PD	689	24.0	24.0	24.000	.0000
08	**PD During School Year**	1007	.000	20.000	12.37550	8.798581
	Summer PD	1007	21.0	24.0	23.611	.9607

Appendix A
(cont'd)

Elementary Summer 2012 PD

Grade Level		N	Minimum	Maximum	Mean	Std. Deviation
03	**PD During School Year**	15	5	20	16.00	6.325
	Summer PD	15	54	60	59.20	2.111
04	**PD During School Year**	695	5	25	13.63	6.883
	Summer PD	695	54	60	59.07	1.993
05	**PD During School Year**	1293	0	25	13.31	8.247
	Summer PD	1293	24	60	56.34	7.931

Secondary 2012 Summer PD

Grade Level		N	Minimum	Maximum	Mean	Std. Deviation
06	**PD During School Year**	1170	.000	15.000	6.85011	5.716591
	Summer PD	1170	12.0	24.0	22.436	3.3273
07	**PD During School Year**	689	.000	20.000	11.50943	6.283941
	Summer PD	689	24.0	24.0	24.000	.0000
08	**PD During School Year**	1007	.000	20.000	12.37550	8.798581
	Summer PD	1007	21.0	24.0	23.611	.9607

Appendix B

Propensity Score Matched Grade 5 Student Group by Demographic Characteristics for 2012 BCM Teachers		
5th grade	**Students of BCM Teacher Sample (n = 2152)**	**Matched Control Student Sample (n = 2381)**
Covariates	***%***	***%***
Gender		
Male	50.6	50.7
Female	49.4	49.3
Eco (Free Lunch)	36.3	33.6
At Risk	77.6	78.5
G/T	1.5	1.2

Appendix C

Propensity Score Matched Grade 8 Student Group by Demographic Characteristics for 2012 BCM Teachers

8th grade	**Students of BCM Teacher Sample (n = 589)**	**Matched Control Student Sample (n = 1369)**
Covariates	***%***	***%***
Gender		
Male	46.9	51.7
Female	53.1	48.3
Eco (Free Lunch)	31.9	32.0
At Risk	77.6	78.5
G/T	1.5	1.2