Partnering to Improve Mathematics and Science Attainment by Improving Teacher Knowledge: A Case Study of the California Mathematics and Science Professional Development Partnership (CaMSP)

Andrew Thomas, Ph.D.
True Analytics
Los Angeles, CA
athomas@eddetective.com

Patricia O'Driscoll
Public *Works,* Inc.
Pasadena, CA
patty@publicworksinc.org

Mikala Rahn, Ph.D.
Public *Works,* Inc.
Pasadena, CA
patty@publicworksinc.org

Descriptors

Professional development, mathematics education, science education

Abstract

This paper reports select findings from a comprehensive multi-year mixed methods case study and impact evaluation of California's Math and Science Partnership (CaMSP). As the state with the largest MSP allocation, the most partnerships, and a high degree of diversity in terms of LEA size and type, university participation and student demographics, California is a rich case that warrants attention. Utilizing Desimone's (2009) conceptual framework for improvement of professional development impact studies, we offer key insights into how to use the LEA-University partnership model to improve mathematics and science teaching. In addition, we use value-added analysis to show CaMSP is associated with improved student achievement in general mathematics and, in particular, Algebra I. Higher science test scores may also be associated with CaMSP professional development.

Partnering to Improve Mathematics and Science Attainment by Improving Teacher Knowledge: A Case Study of the California Mathematics and Science Professional Development Partnership (CaMSP)

Improving the mathematics and science skills of our nation's schoolchildren has been a policy priority for many years but our knowledge about how best to achieve this goal is still partial, inexact and hampered by the enormous variety of method and quality found in professional development impact studies (Desimone, 2009). This study takes advantage of extensive multi-year data on the implementation and results of the nation's largest cohesive and structured mathematics and science professional development initiative (U.S. Department of Education, 2009), to make key inquiries into the nature and effects of university-LEA partnerships for professional development.

Over a decade ago, calls from scientific, business and government leaders (e.g., Greenspan, 2000) led to the creation of the Improving Teacher Quality grant programs in the area of mathematics and science (ESEA Title II, Part B). This policy initiative, an important part of No Child Left Behind (ESEA, 2001), authorized state education agencies to administer statewide Mathematics and Science Partnership (MSP) competitive grant programs. Underlying these programs is the proposition that partnerships between the mathematics, science, or engineering faculty of colleges or universities and high-need local education agencies (LEAs) are well-suited to provide professional development (American Institute for Research, 1999).

Since the Federal MSP program was created and funds were allocated, MSPs have been established in all fifty states, the District of Columbia and Puerto Rico. All of these partnerships have undergone extensive evaluation and results are promising

(American Institute for Research, 1999; Directorate for Education and Human Resources, 2010). Still, more research is needed to better connect professional development practices to student achievement outcomes and identify promising practices.

Conceptual Framework

Agreeing with Desimone (2009, p. 184) that, “sharing a conceptual framework that defines important features of teacher learning experiences has the potential to move the field forward in terms of building a consistent knowledge base,” the authors used Desimone’s framework to guide this study. The framework has two parts. First, it identifies a set of five core features, or critical dimensions by which professional development can be defined, described or gauged. Second, it proposes a theory of action that describes how programs of professional development function to affect educational outcomes.

The five critical dimensions in the framework are as follows:

- Content focus – a focus on subject matter competence—in the case of MSPs, the professional development is focused on mathematics and science content.
- Active learning – Learning by doing or the extent to which teachers learn through participation in authentic, meaningful experiences.
- Coherence – Consistency with teachers knowledge and beliefs and a connection between what teachers learn in professional development with what they are expected to teach.

- Duration – A substantial threshold for the dosage of professional development that appropriately support the intended student-learning goals.

- Collective participation – taking the form of subject matter teams from the same school or district, professional learning communities, or grade-level teams, among other types of group participation.

The second part of the conceptual framework, the theory of action, relates each of these critical dimensions to several outcomes, by steps. The steps are as follows: the activities of a professional development program affect teacher knowledge, skills, attitudes or beliefs. These changes then manifest themselves in modified teacher instructional behavior or practices. Modified instruction then results in improved student learning (see Figure 1).

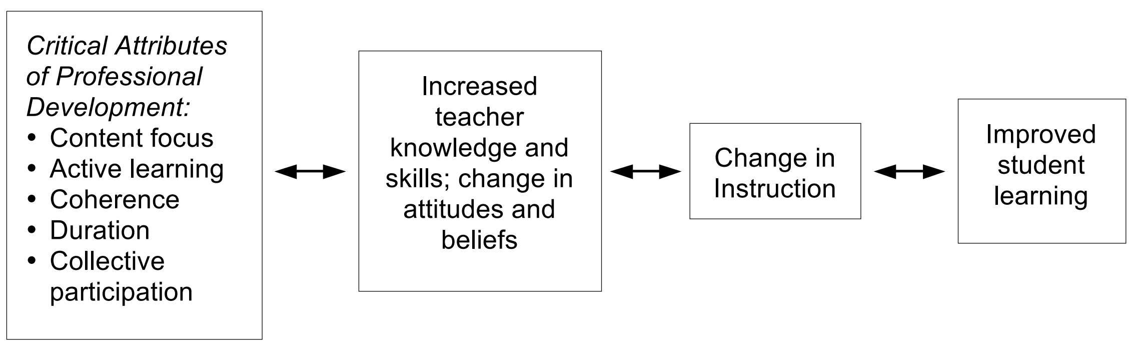


Figure 1: Desimone Conceptual Framework for studying the effects of professional development on teachers and students

Here, we apply the framework to provide categories and a theory of action for a partly descriptive and partly explanatory study of CaMSP.

Methods

Using the critical dimensions of professional development as a guide, we generated the following two research questions. The first question asks whether there is quantitative evidence that CaMSP positively impacted student achievement. The second question focuses on describing characteristics of the CaMSP professional development experiences in terms of the five critical dimensions and asks how these dimensions are related to positive outcomes in individual partnerships.

1. Was CaMSP associated with improvements in mathematics and science student achievement, as measured by the California Standards Test?

2. How and to what extent were the five critical dimensions of professional development manifested in individual CaMSP partnerships that showed evidence of positive student impact? (Note that sub-questions were used to investigate each of the five dimensions. For example, sub-question 1 was: *how and to what extent was a content focus manifested in successful partnerships?)*

Research Context & Sample

Each CaMSP partnership was led by a Local Education Agency (LEA) or multiple education agencies that met the state's requirement to have a student population qualifying for the National School Lunch Program (NSLP) at a rate of 40% or higher. These LEAs were considered "high-need." Lead LEAs formed partnerships with other districts and Institutions of Higher Education (IHEs) based on a variety of factors related to the local and regional context. Partnerships were required to include science and mathematics departments at universities. They may have included associations with education departments as well.

For the current study, CaMSP involved 88 different partnerships between local education agencies (LEAs) and Institutions of Higher Education (IHEs).[1] Of these partnerships, 54 were mathematics-focused, 21 focused on science, and another 13

[1] The lead LEA in a few partnerships is a County Office of Education. In these cases, the COE is required to serve teachers from its own programs as well as teachers from partner LEAs.

provided both mathematics and science professional development.[2] As a consequence of the California Department of Education's funding cycles, partnerships were clustered into five cohorts of teachers, each of which began professional development activities at a different time. Each cohort involved between six and nineteen LEAs. The mathematics partnerships served grade levels three through Algebra I, which is taught in both middle school and high school. Science partnerships served the third through the eighth grades.[3]

In the 2007-08, 2008-09 and 2009-10 school years, PW collected teacher rosters from all partnerships and matched teachers to California student data. After matching partnership treatment teachers to similar non-partnership comparison teachers (as described below), PW consolidated complete data for the 2007-08 school year on 227,887 mathematics students and 56,651 science students, taught by 1,581 treatment teachers and 3,669 similar comparison teachers from throughout the state. For the 2008-09 school year, PW consolidated complete data on 120,535 mathematics students and 44,674 science students, taught by 1,055 mathematics and 539 science treatment teachers, and 1,826 mathematics and 422 science comparison teachers. In 2009-10, PW consolidated data for 116,291 mathematics students and 20,040 science students taught by 1,146 mathematics and 528 science treatment teachers, and 1,826 mathematics and 822 science comparison teachers.

Data sources and collection procedures

Qualitative and quantitative data was collected using partnership site visits and professional development observations, two statewide surveys—one of participating

---

[2] Partnerships in Cohort 4 and after had to select mathematics *or* science and could not serve both content areas.

[3] Prior to 2009, partnerships could only serve fifth grade through Algebra I for mathematics and fourth through eighth grade in science.

partners and another of teachers—telephone interviews, teacher personnel data, and student demographic and achievement data.

Matching procedure

A quasi-experimental design was required to answer research question 1, which asked about program impacts on student achievement. Since CaMSP was not a randomized trial, we constructed a two-level matched comparison group. First we matched non-participating teachers to participating teachers in terms of: partnership, school district, school, grade level, years teaching and credential level. Once matching teachers were found, academic performance and demographic data was collected for students of both treatment and comparison teachers, producing a database of over 280,000 students. Students from the two groups were then matched using a software procedure called "Coarsened Exact Matching," or CEM (Iacus, King, & Porro, 2008), which found one-to-one exact matches—or virtual twins—for all of the treatment students. After matching, statistical tests of difference were conducted and, where appropriate, OLS regression estimates of CaMSP effects on a range of dependent variables were made.

Findings

Was CaMSP associated with improvements in mathematics and science student achievement, as measured by the California Standards Test?

Several analyses were conducted with the matched treatment and comparison groups and are presented here: pooled analyses with all mathematics cohorts for all three years, a partnership analysis within cohorts for 2009-10, and a separate set of science partnership analyses for all three year.

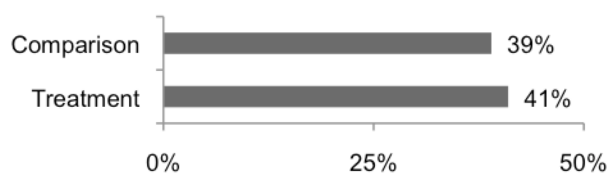
Combined Mathematics Pooled Analysis

Since CaMSP has affected thousands of students in the six years of its implementation in California, we asked whether the overall educational attainment of the California mathematics students who have been taught by participating teachers was higher compared to students who have not been taught by participating teachers. To explore overall mathematics achievement, PW conducted separate pooled analyses that examined mathematics proficiency levels and scaled scores across all cohorts and partnerships for the school years 2007-08, 2008-09, and 2009-10.

Matching virtual twins exactly on a series of covariates meant that the empirical distributions of background variables for both groups were identical and therefore unrelated to the treatment variable (having a partnership teacher). Balancing the groups (treatment and comparison) in this way made further controlling for observed covariates unnecessary. Simple tests of difference were adequate to estimate causal effects (average effect of the treatment on the treated—ATT). We used χ^2 (Chi-squared) and t-tests to examine differences in proficiency levels and differences in scaled scores, respectively, between the treatment and comparison groups.

However, we also wanted to know the magnitude of the impact CaMSP had on mathematics attainment. Was it more or less influential than a student's background and prior achievement? To answer these questions, when a preliminary analysis indicated an effect, we used OLS regression to estimate the effect size and effect direction of the treatment, holding demographic and prior achievement variables constant at their means.

Results from the 2008 pooled analysis showed first, treatment teachers had students who more often performed at or above grade level (proficient and advanced) in mathematics when compared to the matched students of comparison teachers (see Figure 2). Second, the difference in mean combined mathematics CST scaled scores between the two groups was statistically significant, with the treatment students performing better (341.39 v. 339.81, p<.0001). But the difference of less than two scaled score points was so small as to be less than meaningful from a practical standpoint.



(χ2=20.36, p<.0001)

Figure 2. Pooled Sample, Matched Comparison to Treatment, Percent Proficient or Advanced CST Combined Mathematics, 2007-08

In order to determine the relative influence of various factors, including the treatment, on combined mathematics achievement (all grades, all tests) in 2008, a complete regression was run for the pooled sample. The dependent variable was the CST Mathematics scaled score of interest (2008, 2009 or 2010)—which represented different tests depending on the grade level of the student and was a measure of overall achievement of the treatment group. The independent variable of interest was an indicator of whether the student had a partnership-trained teacher. Covariates representing the prior year CST mathematics score as a baseline, ethnicity, language classification, poverty (national school lunch eligibility), gifted and talented classification, and special education designation. These variables are known to be associated with student performance on mathematics tests. In order to compare effect sizes, fully standardized beta coefficients were calculated (Long & Freese, 2003).

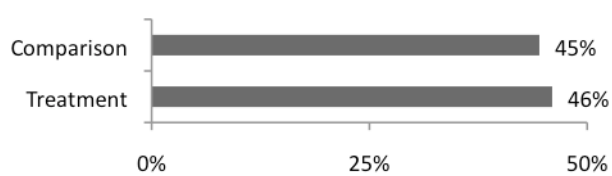
Results were that the treatment – and the model with additional variables added – was significant (see Table 1) and accounted for 60% of the variability in CST mathematics scores (adjusted R^2 .6051). However, the influence of being taught by partnership teachers on CST mathematics performance overall was small. The biggest effect in the model came from prior achievement (β=.71): if a student's performance on the 2007 went up by one standard deviation, his or her performance on the 2008 test was expected to increase by 7/10ths of a standard deviation, provided that all the other variables in the model are held constant. The second most influential factor was gifted and talented status, followed by white and Asian ethnic classification, then poverty, then Hispanic and "reclassified fluent English proficient," the last three of which had negative effects. Other negative factors such as status as an English language learner and special education designation came next in the ranks of influential variables. Last in the list of effects is the treatment—having a partnership teacher, (β=.010).

Table 1.

Regression Analysis of 2008 Combined Mathematics CSTs and Algebra I CST Scaled Scores—with Treatment, Prior Achievement and Demographic Variables as Predictors

VARIABLES	(Pooled) Combined Mathematics β	*t*	*p*	(Pooled) Algebra I β	*t*	*p*
Mathematics Partnership Treatment	0.00980	4.583	0.000	0.0270	5.610	0.000
Prior Achievement (2007 CST)	0.709	279.600	0.000	0.6820	119.500	0.000
Hispanic	0.0329	8.138	0.000	0.0283	3.214	0.001
White	0.0486	14.650	0.000	0.0420	5.735	0.000
Filipino	0.0203	7.624	0.000	0.0407	6.744	0.000
Asian	0.0471	14.250	0.000	0.0644	8.407	0.000
English Only	-0.0174	-6.083	0.000	-0.0383	-6.070	0.000
English Language Learner	-0.0341	-13.290	0.000	-0.0504	-9.249	0.000
Poverty (NSLP)	-0.0326	-13.410	0.000	-0.0393	-7.301	0.000
Special Education	-0.0190	-8.758	0.000	-0.0166	-3.411	0.001
Gifted and Talented	0.0827	34.760	0.000	0.0980	18.320	0.000
Observations	86,190			18,162		
R-squared	0.606			0.584		

Results from the 2009 pooled analysis also showed overall that treatment teachers had students who significantly more often (46%) performed at or above grade level (proficient and advanced) in mathematics when compared to the matched students of comparison teachers (45%, see Figure 3). The difference in mean 2009 Mathematics CST scaled scores between the two groups was statistically significant, with the treatment students performing better (350.63 v. 348.16, $p \leq .001$). But, as in the previous year, the difference between the two groups was very small – less than two scaled score points.



($\chi 2$=19.26, d.f.=1, N=87,732, $p \leq .001$)

Figure 3. Pooled Sample, Matched Comparison to Treatment, Percent Proficient or Advanced CST Combined Mathematics, 2008-09

To investigate how treatment and comparison students made progress toward state goals of proficiency during the year they experienced a CaMSP teacher, we also conducted growth analyses for 2009 student data. Because CST scaled scores are not scaled to be compared year-to-year, we transformed the scores into percentile ranks for all of the students in the dataset, calculating ranks within grade levels. In 2009, treatment students performed at the 51st percentile, moving up slightly from the 50th percentile in 2008, a significant change ($p \leq .001$). Comparison students performed at the 48th percentile in both years, showing no appreciable growth.

Regression for the 2009 pooled analysis, predicting combined mathematics tests scores, showed that the treatment – and the model with additional variables added – was significant and, as in 2008, accounted for 61% of the variability in CST mathematics scores (adjusted R^2 .61). The effect of being taught by partnership teachers was small. As in the 2008 analysis, prior achievement was the most important factor in the model (β=.71): for every additional increase by one standard deviation of a student's 2008 scaled score, his or her performance on the 2009 test was expected to increase by 7/10ths of a standard deviation, when all the other variables in the model were held constant at their means. The next most influential factors were Asian ethnicity and gifted and talented status, followed by white and "reclassified fluent English proficient." Poverty, special education designation and African American ethnicity were negative factors. The treatment—having a partnership teacher—exerted a slight, but significant positive effect (β=.02). See Table 2.

Table 2.

Regression Analysis of 2009 Combined Mathematics CSTs and Algebra I CST Scaled Scores—with Treatment, Prior Achievement and Demographic Variables as Predictors

VARIABLES	(Pooled) Combined Mathematics β	*t*	*p*	(Pooled) Algebra I β	*t*	*p*
Mathematics Partnership Treatment	0.0153	7.264	0.000	0.0376	7.575	0.000
Prior Achievement (2008 CST)	0.0724	282.60	0.000	0.6590	111.500	0.000
Asian	0.0090	31.78	0.000	0.0845	16.130	0.000
Filipino	-0.0186	4.15	0.000	0.0197	3.892	0.000
Black	0.0358	-8.06	0.000	-0.0124	-2.249	0.025
White	-0.0066	12.67	0.000	0.0393	5.855	0.000
Poverty (NSLP)	0.0210	-2.71	0.007	-0.0001	-0.024	0.981
English Only	0.0134	6.25	0.000	0.0266	3.113	0.002
Initially Fluent English	0.0252	5.57	0.000	0.0207	3.513	0.000
Reclassified Fluent English	-0.0099	9.24	0.000	0.0562	7.994	0.000
Special Education	0.0634	-4.63	0.000	-0.0282	-5.546	0.000
Gifted and Talented	0.0724	27.16	0.000	0.0762	14.000	0.000
Observations	87,732			18,208		
R-squared	0.612			0.551		

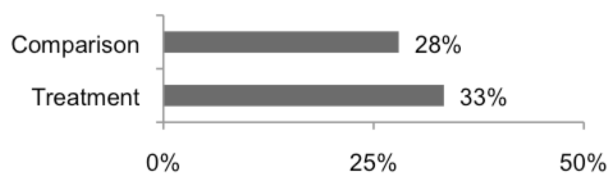
Pooled results for 2010, however, showed a different outcome. While the overall percentage of California school children that performed at or above grade level increased for both treatment and comparison groups, the difference between groups was statistically indistinguishable (49% Comparison, 48% Treatment). On average, both groups achieved CST combined mathematics scaled scores of approximately 350. Therefore, no regression was conducted for 2010 test scores.

The 2010 growth analysis revealed little difference between treatment and comparison groups. In 2010, matched treatment students performed on average at the 50th percentile for their grade, where they had placed also the previous year (2009). Comparison students in 2010 performed at the 48th percentile, which is also where they had placed the prior year. Thus, the 2010 students, whether in the treatment or control group, showed no gain from their achievement in 2009.

Algebra Pooled Analysis

Algebra I scaled scores are of special interest because CaMSP professional development curriculum focuses on strategies for teaching algebra skills more effectively. In order to explore algebra performance, it was important to match only students who took algebra during the 2007-08, 2008-09 and the 2009-10 school years.

In 2008, 33% of the algebra-matched treatment group achieved at grade level (proficient or advanced) on the algebra CST, whereas only 28% of the comparison students attained this level of performance. The average scaled score of treatment twins in 2008 was 330, whereas comparison twins scored only 322 on average (p<.0001).

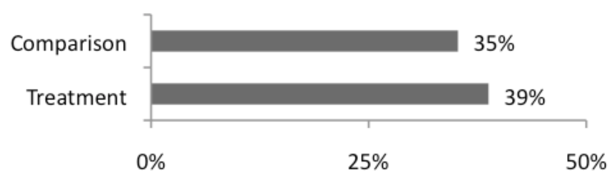


(χ2=61.02, P<.0001)

Figure 4: Pooled Sample, Matched Comparison to Treatment, Percent Proficient or Advanced CST Algebra I, 2007-08

A regression using 2008 Algebra I scores as the dependent variable, taking background and prior achievement factors into account, showed that CaMSP was more influential on Algebra I than it was on mathematics test scores in general (see Table 1). In this model, prior achievement was again the strongest relative predictor of test scores, followed by GATE, and being Asian, Filipino or white. CaMSP was the factor with the sixth strongest relative impact on Algebra I scores (β=.037). In other words, CaMSP was associated with almost four-tenths of a standard deviation increase in mathematics CST scores in 2008.

In 2009, the average algebra score of the treatment twin was 337, whereas the comparison twin scored only 331 on average—a significant difference ($p \leq .001$). In terms of proficiency bands, 39% of the treatment group achieved at grade level (proficient or advanced), whereas only 35% of the comparison students attained this level of performance (see Figure 5).



(χ2=24.12, d.f.=1, n=18,208, p≤.001)

Figure 5: Pooled Sample, Matched Comparison to Treatment, Percent Proficient or Advanced CST Algebra I, 2008-09

2009-10 Partnership Analysis.

To understand the difference between treatment and comparison groups in the pooled analyses of CaMSP, we examined achievement within the individual partnerships that comprise the pooled population for 2009-10, the latest year of CaMSP implementation for which complete data is available. Following the conceptual framework described above, the dimension of professional development that lends itself to quantitative analysis is duration, or dosage. In order to explore the impact of duration of professional development activities on student achievement, we chose to divide the partnerships into three groups based on whether they were at the beginning of their three-year funding period, midway through, or had completed by Fall 2010. We analyzed CST scaled scores and year-to-year gain in CST percentile rank for each of the 41 partnerships active in the 2008-09 and 2009-10 school years.

Beginning partnerships.

Ten beginning partnerships were included in the analysis. All of the beginning partnerships belonged to Cohort 6, which began professional development activities January 2009. By the end of data collection for this study in Spring 2010, these beginning partnerships had completed one long round of funding and one cycle of professional development (60 hours of intensive training and 24 hours of classroom follow up over 18 months). In this time, they produced one year of student test results.

We analyzed two achievement measures for beginning partnerships. They were the CST combined mathematics (multiple tests) average scaled score for matched treatment students in each partnership compared to matched comparison (non-partnership teachers) students in the same year (2009-10), and CST combined mathematics percentile rank gain, comparing how much matched treatment students improved in terms of their performance rank to how much matched comparison students improved in their rank from the year prior to treatment (2008-09) to the year of treatment (2009-10).

Results were mixed. Of the ten partnerships, four partnerships (lead LEAs: Antioch, Central, Fresno and Los Angeles) had average student CST scaled scores in 2009-10 that were significantly better compared to the 2009-10 scaled scores of their matched comparison students and three showed significantly more percentile rank gain between 2008-09 and 2009-10 compared to their virtual twins. One partnership had significantly lower CST scaled scores and lower percentile rank growth measured up against its comparison group.

Midstream partnerships

Ten partnerships had completed two cycles (120 hours of intensive training and 48 hours of classroom follow up) by the time that CST data was collected for this paper in Spring 2010. These partnerships belonged to Cohort 5 and the Research Cohort, both of which began professional development activities in 2008 (Cohort 5 began in January and the Research Cohort began in June). Since these midstream partnerships affected two cohorts of students, we used two years of student test results (2008-09 and 2009-10) to compare midstream partnerships to their matched virtual twins. Thus, four achievement measures were used for midstream partnership analysis: the CST combined mathematics average scaled score for matched treatment students compared to matched comparison students in the same year—2008-09, the CST combined mathematics average scaled score comparison for 2009-10, the CST combined mathematics (multiple tests) percentile rank gain, comparing how much matched treatment students improved in rank to how much matched comparison students improved in their rank from the year prior to treatment (2007-08) to the year of treatment (2008-09); and the CST combined mathematics percentile rank gain from the first year of treatment (2008-09) to the second year of treatment (2009-10). The results of this last indicator are displayed in Table 3.

Two partnerships (lead LEAs Shasta County, Alameda City) showed positive results on all four indicators: higher same-year CST scaled scores, comparing treatment to comparison students, in both 2009 and 2010; and higher two-year growth in average percentile rank between both 2007-08 and 2008-09 and between 2008-09 and 2009-10. Another five partnerships in this category showed a mix of positive outcomes and non-significant outcomes. One partnership (lead LEA: Placer County) appeared to improve from its first to its second cycle, performing less well on its 2009 CST scaled scores and 2008 to 2009 percentile rank growth when compared to its comparison group, but showing significant gain between 2009 and 2010. Two final partnerships had results that did not differ significantly from their comparison groups.

Table 3.

Midstream Partnerships, Matched Treatment vs. Comparison, CST Combined Mathematics Percentile Rank Gain between the 2008-09 and 2009-10 School Years

			Treatment Percentile Rank			Comparison Percentile Rank		
Lead LEA	**Partnership name**	*N*	2008-09	2009-10	Gain	2008-09	2009-10	Gain
Alameda	Strategic and Intensive Mathematics Initiative	404	45th	48th	3%	44th	44th	0%*
Chico	Mathematics Professional Learning Community (MPLC)	466	41st	42nd	1%	41st	42nd	1%
Del Norte	Wild Rivers Math Academy	650	45th	46th	1%	45th	48th	3%
El Rancho	Project Algebra Preparedness for High Achievement	1379	44th	45th	1%	44th	45th	1%
Placer COE	Rigorous Instruction in Mathematics Study (RIMS)	1261	55th	53rd	-2%	55th	51st	-4%*
Red Bluff	North State Math Partnership	529	56th	55th	-1%	56th	51st	-5%*
San Francisco	Partners as Resources to Improve Mathematics Education (PRIME)	461	46th	50th	4%	46th	47th	1%*
Sanger	Central Valley Math Project	1584	53rd	52nd	-1%	53rd	52nd	-1%
Shasta COE	Shasta County Math Partnership (SCMP)	1070	49th	51st	2%	49th	49th	0*
Washington	Washington Union California Mathematics and Science Partnership	237	32nd	38th	6%	32nd	35th	3%*

*$p \leq .05$

Complete partnerships

Twenty-one partnerships had completed all three of their funding cycles by the time data was collected for this report. Each of these complete partnerships implemented 80 hours of intensive, plus 24 hours of classroom follow up professional development in their first year and 60 hours of intensive and 24 hours of classroom follow up professional development in their subsequent two cycles, resulting in a total of 272 hours of professional development for each complete partnership.

Eleven of the complete partnerships belonged to Cohort 3, which completed in June 2009. Ten were Cohort 4 partnerships, which completed their activities in June 2010. For the 11 Cohort 3 partnerships, this paper includes analysis of only two outcome measures (because only one year of data was within the scope of this study): same-year comparative CST average scaled score in 2008-09 and final year percentile rank growth from 2007-08 to 2008-09. For the eight Cohort 4 partnerships, this paper includes analysis of three outcome measures: same-year comparative CST average scaled score in 2008-09 and 2009-10, and final year percentile rank growth from 2008-09 to 2009-10.

Four of the 11 Cohort 3 partnerships (lead LEAs: Aromas/San Juan, Colusa COE, Healdsburg, Santa Maria-Bonita) showed positive and significant results for both measures reported: same-year CST scaled scores and percentile rank gain. One partnership showed negative results, and the remaining six showed no significant difference between treatment and comparison groups on any measure.

Of the complete Cohort 4 partnerships, only one showed positive results on more than two measures: Lincoln Achievement in Mathematics Partnership performed higher on average in terms of CST scaled score—treatment vs. comparison—in 2009. It also showed greater percentile rank growth in 2009 and in 2010 when treatment students were compared to matched comparison students within the partnership. (Table 4 shows the 2009-2010 gain as an illustration of the complete partnership analysis). Another partnership (lead LEA Santa Clara) showed higher relative percentile rank growth two years in a row—2009 and 2010. One partnership showed higher one-year outcomes in 2009 and one showed higher one-year outcomes in 2010. Three additional partnerships showed no significant results in 2009, but negative relative growth in percentile ranks in 2010. A final partnership had non-significant outcomes on all four measures.

Table 4.

Complete Partnerships, Matched Treatment vs. Comparison, CST Combined Mathematics Percentile Rank Gain between the 2008-09 and 2009-10 School Years

			Treatment Percentile Rank			Comparison Percentile Rank		
Lead LEA	**Partnership name**	N	2008-09	2009-10	Gain	2008-09	2009-10	Gain
Alum Rock	South Bay Mathematics Collaborative	1359	42nd	41st	-1%	42nd	46th	4%**
Imperial	Imperial Cty MP	878	48th	47th	-1%	48th	48th	0%
Lincoln	Lincoln Achievement in Mathematics Partnership	708	52nd	54th	2%	52nd	50th	-2%**
Little Lake	Achievement in Little Lake for Mathematics (ALL for Math)	134	52nd	48th	-4%	51st	50th	-1%
Pajaro Valley	Pajaro Valley USD MP	685	43rd	45th	2%	43rd	43rd	0
Pasadena	Pasadena Math Pipeline	987	45th	38th	-7%	45th	45th	0**
Santa Clara	ELDI-ARISE	1318	50th	51st	1%	50th	49th	-1%**
Westminster	Developing Communities of Mathematical Inquiry (DCMI)	1639	57th	57th	0	57th	59th	2%*

*$p \leq .05$, **$p \leq .001$

Partnership summary

The partnership analysis includes up to two years of data for each partnership and was designed to shed light on some of the differences between partnerships that were at different stages of implementation. Beginning partnerships had just completed their first full cycle of professional development activities at the time of data collection for this report (Spring 2010). These ten partnerships – all part of the sixth cohort of CaMSP participants – showed some signs of success. After one-year and up to 84 hours of targeted professional development for each teacher, three of the ten (33%) produced better 2010 CST mathematics test gains compared to their matched comparison group. If more professional development results in greater test gains, it would be expected that midstream partnerships, which had provided teachers with 168 hours of learning support, would show evidence of increased success compared to first-year partnerships.

There were ten midstream partnerships in this analysis, which entailed exploration of both first and second year test results. As with the beginning partnerships, thirty percent of the midstream partnerships produced higher first-year CST mathematics gains. But in their second year, 60% of this group had gains higher than their comparison groups. This finding supports the proposition that more hours of professional development may be associated with greater student achievement.

However, the evidence from the complete partnership analysis did not corroborate the proposition. There were 21 partnerships that had completed three cycles – up to 272 hours of professional development. This analysis explored the second and final years of test score data for some of these partnerships (Cohort 4, ending in 2010) and only the final year of test score data for the rest (Cohort 3, ending in 2009). Overall, 33% of the partnerships experienced more test score gain in their final year compared to their matched comparison groups. However, this result masks the differences between cohorts. Forty-five percent of Cohort 3 partnerships showed more percentile rank gain than their comparison groups, whereas only 25% of Cohort 4 partnerships posted similar results.

Overall, this analysis did not provide clear evidence of a positive trend in CST improvement correlated to increased quantity of professional development. Factors specific to cohorts and individual partnerships probably played greater roles than hours of professional development alone.

This partnership analysis, however, sheds some light on why 2010 outcomes flattened for the pooled CaMSP analysis. Sample size was a factor. 2008-09 data was available for 31 partnerships, 21 of which were complete. Of these 31 partnerships, 17 showed evidence of positive effect, six showed evidence of negative effect and eight showed no significant effect.

2009-10 data, on the other hand, included only 28 partnerships, eight of which were complete. Of the 28, 16 showed evidence of positive effect, six showed evidence of negative effect, and six showed no significant effect. These small differences in sample sizes from year to year, combined with the generally small CaMSP effect size, may be enough to account for the disappearance of significant treatment/comparison variability in 2010 mathematics achievement.

Science Analysis

Evaluation of the science partnerships followed a simpler procedure compared to the mathematics analysis. Because science tests are administered only in fifth and eighth grade, no year-to-year comparisons were possible. Nor was it possible to match virtual twins based on prior performance on a science test. Therefore, this analysis includes only one-year scores (2008, 2009 and 2010), comparing treatment and comparison groups that were matched on background variables and prior mathematics CST proficiency levels. Results were disaggregated by science test (either fifth or eighth grade science). Both pooled and partnership analyses were conducted, though this paper presents pooled analyses only. In addition, regression models were computed when significant differences between groups appeared in the comparisons of means.

Science pooled analysis

The pooled analysis for both grade levels and all cohorts showed a positive effect on combined fifth and eighth grade science scores in 2008. The comparison group had an average scaled score of 339, compared to the treatment group, which scored 343 on average (p<.0001). The proportions of students at grade level (proficient and advanced) were 43% for comparison and 46% for treatment. Students who had taken fifth of eighth grade science from a partnership teacher clearly performed better compared to those students who had not had a partnership-trained science teacher.

A look at grade level differences revealed that higher performance primarily occurred at the fifth grade level. Treatment fifth graders scored almost seven points higher than comparison fifth graders (342 vs. 335, p<.0001). This higher achievement is apparent in proficiency bands as well. Forty percent of comparison students and 44% of treatment students performed at grade level. Eighth graders in the treatment group, however, performed no better than their virtual comparison twins.

To determine effect size for 2008 science scores, two regressions were computed, one to predict the overall science achievement, and one that focused solely on fifth grade performance. Each regression used an indicator of whether students had a science treatment teacher as the independent variable of interest and included special education designation, GATE designation, ethnicity and language classification as the background covariates.

The first model was significant and accounted for 34% of the variability in combined science test scores. Having a science partnership teacher produced a beta coefficient of .019. The model that predicted fifth grade performance had a larger positive effect (β=.039) and accounted for 35% of the test score variability.

Table 5.

Regression Analysis of 2008 Combined Science CSTs and Fifth Grade Science CST Scaled Scores—with Treatment, Prior Achievement and Demographic Variables as Predictors

VARIABLES	(Pooled) Combined (5^{th} & 8^{th} Gr.) Science CSTs β	*t*	*p*	(Pooled) 5^{th} Grade Science CSTs β	*t*	*p*
Science Partnership Treatment	0.0191	3.548	0.000	0.0277	3.578	0.000
White	0.119	18.84	0.000	0.124	13.58	0.000
Filipino	0.0542	9.887	0.000	0.0368	4.650	0.000
Asian	0.111	20.10	0.000	0.0684	8.548	0.000
African American	-0.0735	-12.43	0.000	-0.0747	-8.994	0.000
English Language Learner	-0.228	-34.29	0.000	-0.247	-26.14	0.000
Redesignated Fluent Eng. Prof.	0.0626	9.670	0.000	0.0563	6.411	0.000
Poverty (NSLP)	-0.0706	-12.02	0.000	-0.118	-13.92	0.000
Special Education	-0.108	-19.75	0.000	-0.0914	-11.78	0.000
Gifted and Talented	0.347	61.67	0.000	0.338	41.93	0.000
Observations	22,828			10,882		
R-squared	0.339			0.355		

The 2008-09 pooled analysis for grade levels, all cohorts and all partnerships showed a slight significant effect of CaMSP on combined fifth and eighth grade science scores in 2009. The comparison group had an average scaled score of 363, compared to the treatment group's average score of 367, which was a significant difference ($p \leq .05$). Comparison of the proportions of students at grade level (proficient and advanced) and students below grade level showed that 55% of the comparison students and 57% of the treatment students were at grade level or above—a significant result ($p \leq .05$).

Comparison of grade level scaled scores revealed that treatment fifth graders performed no better than comparison fifth graders (353 vs. 352, not significant), but that treatment eighth graders did perform significantly better than comparison eighth graders—371 and 367, respectively ($p \leq .05$). This finding that the partnerships had more impact on eighth grade then fifth grade was the opposite of the finding from the previous year.

The 2009-10 pooled analysis for grade levels, all cohorts and all partnerships showed that the comparison group had an average scaled score of 368, compared to the treatment group's average score of 369, which was not a significant difference. However, a coarser comparison—the proportions of students at grade level (proficient and advanced) compared to students below grade level were 56% for comparison and 58% for treatment, and significant at the 10% level.

Comparison of grade level results showed that, in terms of scaled scores, there no significant differences between groups. Treatment fifth graders scored about the same as comparison fifth graders (357 vs. 358, not significant). Eighth graders in the treatment and comparison groups also were almost identical in terms of scaled score: 374 vs. 373. In terms of proficiency level, though treatment fifth graders performed no better than comparison fifth graders, treatment eighth graders did perform significantly better than comparison eighth graders—60% to 57%, respectively ($p \leq .05$).

Science Summary

As with the mathematics pooled analysis, impact of the science partnerships on same-year test scores was found in 2008 and 2009, but was less apparent in 2010. In 2008, fifth graders seemed to benefit more from being taught by partnership-trained teachers, whereas in 2009 and 2010, eighth graders seemed to receive more benefit.

Due to the lack of the data for studying score gain, or change over time for the same students, this impact study of science partnerships is inherently limited. It should be interpreted as evidence that science partnerships show some signs of positive effect. A partnership-by-partnership analysis (not reported here) revealed variability between partnerships: some managed to improve achievement more than others.

How and to what extent were the five critical dimensions of professional development manifested in individual CaMSP partnerships that showed evidence of positive student impact?

After quantitatively analyzing CST achievement data for positive impact, we found considerable evidence that overall, CaMSP had a small, but measurable positive impact on CST scores in combined mathematics (all tests, all grades) and Algebra I tests. We also found that this slight effect was produced as a consequence of some partnerships showing strong evidence of positive impact and other partnerships showing no evidence that they exerted an effect at all. Consequently, we either asked what dimensions of professional development (following Desimone's framework) were present in the strong partnerships that may not have been present in the other partnerships, or how these dimensions were manifested differently in strong partnerships than they may have been manifested in other partnerships.

In this section, we describe how the five dimensions were manifested in CaMSP in general and then present some preliminary insights about what dimensions matter the most in terms of explaining the variable impact from partnership to partnership. As of Spring 2011, final data collection (phone interviews) is being conducted to gather evidence that would either confirm or contradict these insights.

Content Focus

Content focus was a dimension of relative uniformity across all partnerships in this study. CaMSP partnerships currently select a content focus of either mathematics or science. In previous cohorts, partnerships could provide training in both content areas, though these typically involved different cohorts of teachers and separate training. Thus, CaMSP models are now focused on one or the other content area, and there is little integration across the two content areas. In each cohort, about a third of the partnerships target science and the remaining partnerships target mathematics. In a few partnerships, lessons learned from implementation of what worked in mathematics or science have been used to design proposals for funding of the other content area under a new cohort.

Professional development models in CaMSP incorporate instructional strategies that blend both conceptual understanding and hands-on learning with the need for students to have factual knowledge and computational fluency at their fingertips across both subject areas. Though observations of CaMSP professional development indicated training that was weighted heavily toward teacher understanding of challenging content and pedagogical content knowledge beyond the particular grade levels taught by the teachers, in terms of content decisions and alignment to standards, textbooks, and accountability measures, site visits and interviews indicated that the current framework for accountability in California is a significant factor guiding decisions about what to cover in training.

Site visit observations of professional development activities and analysis of documents and handouts collected during these activities showed that all of the professional development offered through CaMSP was aligned to academic content standards in mathematics and science. District adoption of state-approved textbooks and instructional materials in science and mathematics was also aligned to training in many partnerships, especially those involving single districts. However, in multiple district partnerships, there were multiple adopted texts. Observations of professional development activities noted that guidance on the use of adopted textbooks was often reserved for the time set aside for team planning and collaboration, in coaching sessions, and during lesson planning sessions.

Active Learning

CaMSP professional development consistently modeled the kinds of instructional strategies that blend both conceptual understanding and hands-on learning along with the factual knowledge and computational fluency that both teachers and students need in mathematics and science. In each year of the evaluation, observations of professional development and interviews with trainers, coaches, and others indicate that maintaining teacher interest is of utmost importance—both to encourage attendance and to ensure that each year teachers have an incentive to continue. This plays out in the professional development offered in a variety of ways. For instance, some partnerships have developed on-line resources and communication tools as a strategy to maintain interest and to help with classroom planning. Others have spent more time with science notebooks as a strategy to bolster both student science learning and writing, which is an important component of English language arts accountability measures.

Research in the field of professional development has indicated that teachers need not only high quality information during trainings, they need opportunities to practice and to plan how to use what they have learned. Classroom follow-up hours are intended to answer the frequently heard concern of teachers that they do not have time to plan how to implement what they have learned and are the component of CaMSP in which the concepts of both active learning and coherence can be manifested most but which also vary the most from partnership to partnership.

For current CaMSP partnerships, classroom follow-up hours are achieved in a variety of ways—some partnerships use coaching models, others have implemented lesson study, some a combination of the two, and still others emphasize collaborative planning time for teacher teams that are not necessarily facilitated by trainers, coaches, or district staff.

In theory, coaching, lesson study, and professional learning communities are elements of a continuum of professional collaboration. Coaching emphasizes supporting individual teachers in their classrooms. Lesson study involves the facilitation of teams of teachers to improve practice based on an action-research model first developed in Japan. Professional learning communities take on a school wide approach whose aim is to embed ongoing professional development in the organization and decision-making of the school. These collaborative structures are organized in many ways and can involve district teachers on special assignment, content coaches, university professors or trainers from professional development providers, but generally emphasize ways to incorporate content, assessment, and teaching strategies.

The classroom follow-up component continues to have the least consistency in terms of quality and organization, though there has been improvement since the beginning of CaMSP. In the first several cohorts, classroom follow-up was uneven and many partnerships had a difficult time getting teachers to complete the hours.

Coherence

Coherence between the professional development delivered and what participants are expected to teach was embedded in CaMSP through the structure of required intensive hours and classroom follow-up organized through coaching, lesson study, or professional learning communities. There is evidence of professional development that was more tailored to both the individual learning needs of participants and the systemic needs of LEA partners related to standards-based instructional materials, pacing guides, and district assessment systems.

Teachers participating in CaMSP report enthusiasm about the infusion of professional development that is engaging and challenging. However, partnerships report that the teachers who remain for the duration of training typically do so because the support they receive throughout the year is adapted by the partnership team to be as tailored to their needs in the classroom as possible.

As CDE has clarified its expectations for the classroom follow-up component, partnerships have focused more on the importance of the classroom follow-up hours and participants have had more success with completion of the hours. In partnerships in which the classroom follow-up has a solid plan and direct connection to the intensive content-focused professional development, teachers have valued the opportunity for collaboration and support, with many expressing appreciation for the coaching or lesson study resources dedicated to the classroom.

For some partnerships, classroom follow-up is refined once the particular needs of the cohort of teachers become clarified. For instance, several partnerships proposed a combination of coaching and lesson study for classroom follow-up in their models. After the summer intensive institutes, many realized that individual teachers had a greater need for coaching and building toward the collaborative and team-based nature of lesson study. The use of lesson study is present in many partnerships, though it is not necessarily implemented consistently with the original design of Japanese lesson study.

The operating definition, implementation design, and the model of lesson study varied from partnership to partnership. In addition, facilitation of lesson study also varied. In some partnerships, lesson study teams were facilitated by district coaches (some trained in lesson study by an IHE partner) or by facilitators with content expertise from IHE partners or by facilitators from an outside provider.

The implementation of the lesson study model of professional development continues to be refined and this strategy is a work in progress—adapted from various approaches to the local context and willingness of teachers to examine and self-reflect about their teaching. In some partnerships where lesson study was new to teachers, the partnership began the process through school-based collaborative teams, often using group and individual coaching strategies as a first step to lesson study. Some partnerships also viewed lesson study as a more sustainable strategy for classroom follow-up because new facilitators are often drawn from participating teachers.

Coaching is also a component of professional development that is offered by many partnerships. Like lesson study, the type of coaching offered varies from partnership to partnership. However, coaching continues to be more defined than other forms of follow-up and partnerships commented about the importance of coaching reflecting a strong relationship between the coach and the participant.

Coaches typically deliver model lessons, observe teachers, hold debriefs about instructional practice, and coach teachers about specific strategies. Most coaches keep logs of their interactions and many partnerships have developed different tools and forms to support the development of the coaching relationship and to have consistency in delivery of classroom follow-up. According to participating teachers, for the relationship to be a success, coaches need to provide valued services to participants and, at their core, be respected classroom teachers. As resources get stretched thin, coaching is often at the mercy of budget cuts, and in districts in which CaMSP resources supplemented existing coaching models, some partnerships have had to adapt by cutting individual coaching hours and working more with collaborative teams or having school-based teams submit logs of what they had accomplished in order to meet the grant requirements. Partnerships are learning that it is important to establish a model for classroom follow-up that can be embedded in teacher practice once the CaMSP funding period is over.

Duration

CaMSP partnerships all developed models for professional development that included 60 hours of intensive professional development supported by 24 hours of classroom follow-up over three years for the same cohort of teachers. Within these general guidelines, partnerships established models that blend research-based strategies with the challenging logistical context of California public schools—including staff reassignments to new grade levels and schools, transitions to other districts, budget cuts, and accountability mandates that do not always support the content areas or other components of the CaMSP professional development model.

Despite these challenges, in general, interviews and site visits indicated that partnerships have become clearer about the distinction between the purpose of intensive professional development and classroom follow-up and have refined how classroom follow-up is organized so that it is more meaningful and beneficial to participants.

In the beginning of CaMSP implementation, many partnerships planned to conduct two-week summer institutes based on lessons learned for quality professional development from successful engagement of teachers through the California Mathematics and Science Subject Matter Projects and other national grants supporting mathematics and science learning such as the Eisenhower initiative and others. Professional development providers knew that engaging teachers in the summer above and beyond what was required by their districts entailed both challenging content and instructional strategies for student engagement that was directly usable when they returned to school. The development of the 80-hour summer institute was further enhanced by the growing consensus that professional development based on the “one-shot” or textbook company format was ineffective but so typical of what teachers generally received during the school year.

However, with CaMSP structured as a partnership between Institutions of Higher Education (IHE) working in collaboration with districts to design a program appropriate for the local context and staffing, partnerships had to shift to a model in which the district was primarily responsible for recruitment and retention of teachers rather than putting on institutes that catered to the interests of individual teachers. While individual teacher interest is important for professional development to be successful, CaMSP partners and professional development providers learned over time to shift to the model of integrated training envisioned by CaMSP in which training resources are invested in cohorts of teachers who receive substantial training over time and have opportunities to embed their learning back in their classrooms and in planning instruction.

CaMSP partnerships, for the most part, rely on IHE trainers, some of whom were experienced with the Summer Institute format and others who were new to professional development but had been brought into CaMSP based on specific content expertise, to develop the content and format of providing the intensive hours in conjunction with district staff who was typically responsible for implementation of the follow-up model. In a few places, IHE partners were involved in both aspects, but this was relatively rare.

As CaMSP has been implemented, the summer institute format has been adapted significantly, with most partnerships offering a one-week institute, sometimes more than once to capture all members of the cohort, or institutes of a few days in length. With the CDE requirement that all members of the cohort complete 30 of their 60 intensive hours prior to August 30th, most of the summer institutes are approximately 30 hours in length. Additional intensive hours are delivered in a variety of ways including an on-going series of one-day or after-school trainings occurring regularly throughout the school year or several one-day workshops delivered on Saturdays or after school. Partnerships have also had participants work during the school year with IHE partners on projects and research related to their summer work during the school year.

In previous cohorts, “makeup” sessions had been offered frequently in order to ensure that hours were completed. However, there was concern that the availability of makeup sessions was hampering efforts to deliver a coherent program of training to all participants. Now there is much less make up offered to participants—rather, project directors work hard to schedule activities around the needs of participating teachers and strengthening the incentives for participation.

Since teachers must complete both a minimum number of hours per year and agree to participate for three years in the program, the teachers participating in professional development have coalesced into solid cohorts with the interest and dedication necessary to participate in multiple years of training. In addition, the stringent requirements for participation have encouraged professional development providers to work closely with LEA partners, resulting in closer ties to district priorities. For teachers, the duration of training has both benefits and challenges. Movement to other grades and schools, layoffs, and natural attrition continue to be difficult challenges to the CaMSP model.

Collective Participation

The strengthening of CaMSP requirements has resulted in cohorts of teacher participants that, over the course of three years of training, are strong in terms of working relationships and teachers with the capacity for leadership in their districts. However, one need identified in the study was a stronger connection to school site administration in the planning and implementation of the classroom follow-up component. From the practical standpoint of integrating CaMSP classroom follow-up with site-based initiatives to avoid conflicts with other priorities, consistent integration of site administrators is an area for growth within CaMSP. In addition, site administrators have been identified as key to sustaining the momentum from individual teacher leaders.

Discussion and Key Insights

The preceding analyses provide insights into the impact of CaMSP on mathematics and science achievement in California in 2008, 2009 and 2010. Some outcomes in 2008 and 2009 were positive. Same-year combined mathematics CST scores comparing the partnership treatment group to the matched comparison group showed positive effects both for all teachers and their students from all cohorts and partnerships pooled together, for individual cohorts in 2008 and for several individual partnerships in 2009. Year-to-year differences in combined mathematics CST scores corroborated the finding from the same-year scores analysis. An examination of Algebra I scores, comparing the treatment and comparison groups in the same year (2008 and 2009) and exploring whether Algebra I performance improved compared to other tests taken the prior year (2007 or 2008), showed that overall the pooled treatment groups performed better on the Algebra I test. However, the investigation of 2010 outcomes did not produce similar results: the pooled treatment group, on average, performed no better than its comparison group, either on combined mathematics or Algebra I.

In addition to analyzing the pooled data from CaMSPs (all cohorts, all partnerships), we examined the role individual partnerships played in overall CaMSP outcomes. We divided partnerships into three groups: *beginning, midstream,* and *complete.* The expectation was that beginning partnerships may show less substantial and significant outcomes, but that student performance would increase as the partnerships completed additional training and implementation developed and was refined over time.

The exploratory analysis completed here provides a way to make coarse comparisons between the three groups.

For the beginning group of partnerships, about 40% showed some signs of positive effect, 10% showed a decline and 50% were not associated with an effect at all. The midstream partnerships—the only group with complete two-year data—looked considerably stronger. There were 10 partnerships in this group. Seventy percent had mixed (non-significant and positive) or consistently positive results, 20% had consistently non-significant results and one partnership improved from 2009 to 2010, initially appearing less effective relative to its comparison group and then posting a higher score.

The complete partnerships present a more complicated picture. These partnerships were from Cohort 3, which completed in 2009, and Cohort 4, which completed in 2010. This study utilized one year of data for analysis of the 11 complete Cohort 3 partnerships and two years of data for the 8 complete Cohort 4 partnerships.

About 40% of the complete Cohort 3 partnerships with one year of data showed mixed positive results, about 10% showed negative results and over 50% did not produce significant results at all. Of the eight complete Cohort 4 partnerships with two years of data, about 40% were mixed positive non-significant, 40% mixed negative non-significant, 10% non-significant on every measure, and one partnership (roughly 10%) that improved from 2009 to 2010. Thus, results from the complete Cohort 4 partnership analysis resembled those from the midstream analysis.

This exploration into the differences between and among partnerships that had just begun, were midway through, or had completed the implementation of their CaMSP grants suggests that implementation details are important in the success of any partnership. It also suggests that to accurately measure effectiveness, it is important to analyze more than one year of data. Some partnerships start out with weaker scores and then move ahead after the first year and very few are consistently effective from the beginning.

The qualitative analysis followed on the impact study and sought evidence of variability between strong partnerships and partnerships that showed little or no effect. What were the hallmarks of the successful partnerships that differentiated them from the rest? To this end, we asked how and to what extent the five critical dimensions of professional development were manifested in individual CaMSP partnerships that showed evidence of positive student impact. Using observational site visits, document analysis and interviews, we examined each partnership in terms of content focus, active learning, coherence, duration and collective participation.

We found little variability in terms of content focus, duration, or collective participation. These three dimensions were broadly defined by the structure of CaMSP itself. A focus on mathematics and science content, of course, was built into the very definition of the program. The California Department of Education required a certain number of hours and a certain level of teacher participation to which each partnership must adhere. These requirements effectively smoothed differences in duration per year and collective participation.

However, we did find considerable variability in the dimensions of active learning and coherence. In terms of both of these dimensions, we largely focused on the classroom follow-up component of CaMSP implementation. Classroom follow-up was generally some combination of coaching, lesson study and collaborative planning time. Preliminary analysis of qualitative data suggests that the appropriateness of this implementation to the local context and needs of teachers as well as the quality of the coaching or lesson study is associated with more positive outcomes and more effective partnerships.

Perhaps more variable and important than the quality of the active learning, however, is the dimension of coherence and how it was manifested partnership to partnership. One of the weak links reported by many, if not most, partnerships in CaMSP, which explicitly prohibits funding of professional development for administrators, was the lack of involvement by school site principals. Though this aspect of implementation, seen as a barrier to effective professional development, varies from partnership to partnership, it is a component that many partnerships are trying to address, particularly those that have had multiple CaMSP grants and continue to see engagement by site administrators in implementation of mathematics and science improvements as a key to long term success. However, from the summer intensives, to the professional development sessions organized during the school year, to the classroom follow-up piece of training, successful implementation of a CaMSP partnerships lends itself to the development of teacher leaders with partner LEAs that have opportunities for growth with the skills they have developed as participants and coaches or facilitators. It is the combination of active and engaged learning in all aspects of professional development with coherence in terms of district policies related to instruction, materials, pacing guides, and the like and the support personnel selected for the partnership that can be seen most clearly in many successful CaMSP partnerships.

References

American Institute for Research. (1999). *Designing Effective Professional Development: Lessons from the Eisenhower Program*. Washington, DC: US Department of Education Office of the Undersecretary, Contract No. EA97001001

Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher, 38*(3), 181-199.

Directorate for Education and Human Resources. (2010). *National Impact Report: Math and Science Partnership Program*. Washington, DC: National Science Foundation

Greenspan, A. (2000, September 21). FRB: Testimony, Greenspan -- The economic importance of improving math-science education. Retrieved July 6, 2010, from http://www.federalreserve.gov/boarddocs/testimony/2000/20000921.htm

Iacus, S. M., King, G., & Porro, G. (2008). Matching for Causal Inference Without Balance Checking [Electronic Version]. Retrieved March 13, 2010, from http://gking.harvard.edu/files/abs/cem-abs.shtml

Long, S. J., & Freese, J. (2003). Estimation Testing, Fit, and Interpretation. In *Regression Models for Categorical Dependent Variables Using STATA*. College Station, TX: Stata Press Publications.

U.S. Department of Education. (2009). MSP Budgets by State [Electronic Version]. Retrieved July 6, 2010, from http://www.ed-msp.net