Using Technology to Enhance the *Art and Science of Teaching* Framework

A Descriptive Case Study

September 2013

MARZANO RESEARCH LABORATORY
www.MarzanoResearch.com
Phone: 888-849-0851 | Fax: 866-801-1447

Marzano Research Laboratory

OUR MISSION
To provide the best research, the most useful actions, and the highest level of services to educators.

OUR VISION
To continuously develop tools that translate high quality educational research into practical applications educators can put to immediate use.

OUR GOAL
To be the place educators go for the latest information and data, synthesized into clear, concise resources that facilitate immediate action.

REPORT AUTHORS

Mark W. Haystead & Anthony J. "Sonny" Magaña

For citation purposes, please refer to this document as:

Haystead, M. W., & Magaña, S. (2013). *Using technology to enhance the* art and science of teaching *framework: A descriptive case study*. Centennial, CO: Marzano Research Laboratory.

BUSINESS DEVELOPMENT OFFICE
555 N. Morton Street
Bloomington, IN 47404
Phone: 888-849-0851
Fax: 866-801-1447

RESEARCH CENTER
9000 E. Nichols Ave. Ste. 112
Centennial, CO 80112
Phone: 303-766-9199
Fax: 303-694-1778

www.MarzanoResearch.com

Table of Contents

Table of Contents i

Introduction 1

Background 2

Method 4

Results 7

Summary and Discussion 20

Technical Notes 26

References 30

Introduction

American education systems have spent billions of dollars on technology tools since the turn of the 21st century (Richtel, 2011). This trend demonstrates an overarching systemic optimism about the potential of technology to solve the various problems that plague modern education. Unfortunately, research on the impact of educational technology on student achievement and engagement is, at best, mixed (Richtel, 2011). The majority of evidence on the influence of technology on student achievement has, by and large, not met the hopeful perceptions of educational technology consumers.

One reason for the mixed influence of educational technology is the provision (or lack of provision) and nature of professional development for teachers. It is an unfortunate reality that, once provided with expensive educational technologies, many teachers do not receive the professional development necessary to make effective use of those technology tools in the classroom. Instead, professional development is often primarily focused on the functionalities of the technology tool itself (for example, computers, software suites, laptops, tablets, interactive whiteboards [IWBs], student response systems [clickers], and so on) and not on how to utilize these tools within the context of an effective instructional framework (such as the framework articulated in *The Art and Science of Teaching*; Marzano, 2007).

Beginning in the 2009–2010 academic year, action research was conducted over a three-year period by teachers in a high-needs, low-achieving elementary school in Southern California. This report presents a descriptive case study of the action research through visual analyses of school proficiency data from statewide student assessments from 2004 through 2012. In addition, this report presents statistical analyses of school and district proficiency data from 2009 and 2012.

The descriptive case study sought to address the following research question:

> Does teacher professional development focused on using interactive technologies to enhance instructional strategies (articulated in the *Art and Science of Teaching* framework; Marzano, 2007) influence student engagement and achievement?

Teachers received 40 hours of in-person professional development, classroom coaching, and online support to improve their understanding and application of interactive technologies (that is, IWBs and clickers) to enhance their utilization of instructional strategies articulated in the *Art and Science of Teaching* framework (Marzano, 2007). It is important to note that the framework provides teachers with a logical planning sequence for effective instructional design.

The IWBs enabled teachers to present academic content to students using multiple sources of media (for example, text, images, sounds, animations, or videos). The clickers enabled all students in a classroom to respond to multiple-choice questions posed by the teacher and allowed teachers to display the polling data from each student in the class.

Although the descriptive case study involved IWBs and clickers, a reasonable inference can be made that the findings may be generalized to any interactive technology with a substantively similar functionality. In other words, the findings may apply to other multimedia technologies that enable teachers to present academic content (such as interactive projectors or desktop computers, laptops, or tablets connected to a projector) and technologies that enable multiple students to respond to multiple types of questions posed by a teacher (such as software or app-based student response systems running on laptops, tablets, or students' smartphones).

Background

During the 2008–2009 and 2009–2010 academic years, Marzano Research Laboratory (MRL) was commissioned by Promethean Ltd. to conduct evaluation studies of the effects on student academic achievement of Promethean ActivClassroom tools and technologies (Marzano & Haystead, 2009; Marzano & Haystead, 2010). IWBs and clickers were among the interactive technologies teachers used in the evaluation studies. It is worth noting that the combined results from the two evaluation studies involved 123 classroom teachers and nearly 5,000 students representing a wide range of schools (urban, suburban, and rural) across the United States, all grade bands (kindergarten through grade twelve), and core content areas (English language arts, mathematics, science, and social studies).

A quasi-experimental research design was used in both evaluation studies. Participating teachers conducted individual experimental-control studies with intact classes of students. In the individual studies, teachers collected pretest and posttest scores from two of their classes of students. One class taught new content with interactive technology (experimental group) and the other taught the same content without interactive technology (control group). Stated differently, in the experimental group, teachers used the interactive technology to enhance their instructional methods, while in the control group, teachers utilized the instructional strategies they would normally employ without using the interactive technology.

Both evaluation studies were conducted in two phases. In the first phase, data were analyzed separately for the individual experimental-control studies. Pretest scores were used as a covariate in an analysis of variance to statistically equate the students and partially control for differing levels of background knowledge and skill. The results of the analyses were used to estimate separate effect sizes for each experimental-control study. It is important to note that the effect sizes were not directly calculated as the standardized difference between the experimental and control group mean scores (averages), commonly referred to as Cohen's *d*. Instead, the effect sizes reflected an adjustment of students' posttest scores based on a common pretest score for the experimental and control groups. The second phase was conducted in two steps. The first step involved a videotape analysis and rating of the instructional methods used by teachers during their classroom instruction with technology and the second step involved an analysis of the relationship between the effect sizes from the first phase and the videotape ratings.

A cautionary note must be made here. As mentioned previously, both evaluation studies employed intact groups (that is, students were not randomly assigned to experimental and control groups). Although analysis of covariance was used to statistically equate students in terms of prior academic achievement based on teacher-designed preassessments, arguments about causal relationships are not nearly as strong as they might be if group membership had been determined through random assignment. However, reasonable inferences can be made about specific findings in each study.

Figure 1 displays the theory of action for both the first-year study and continuation study.

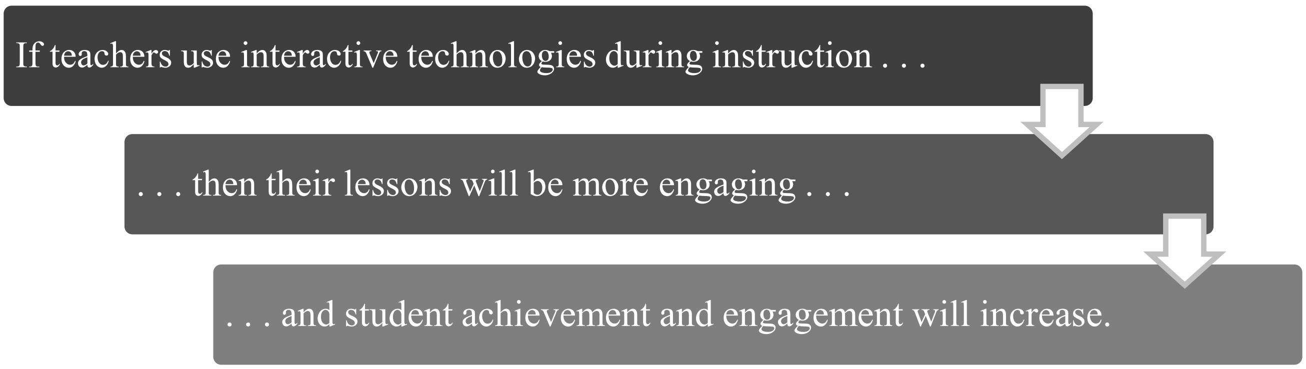

Figure 1: Theory of action for the first-year and continuation evaluation studies.

Each teacher received approximately six hours of introductory technology training prior to participation in the evaluation studies. The training focused on showing teachers how to use the IWB and IWB software to deliver instructional content. Additionally, the training demonstrated how to ask questions using IWB software so that all students could respond using clickers. It is important to note that this introductory training did not address any specific pedagogical elements or instructional strategies, nor did it address how to use the interactive technology to support an instructional framework.

During the first phase of analysis, findings from the first-year and continuation evaluation studies were averaged into a weighted estimate of effect size, $d = 0.36$, 95% CI [0.25, 0.46], $k = 131$. The estimated summary effect size suggested that, on average, students learning new content from teachers using the interactive technology (experimental group) would be expected to outperform students learning the same content without the technology (control group) by 14 percentile points (Marzano & Haystead, 2010).

In 77% of the individual experimental-control studies, students learning new content with the technology (experimental group) outperformed students learning the same content without the technology (control group). In other words, the interactive technology had a positive effect on student achievement as measured by the differences in adjusted assessment scores between the experimental and control groups. Conversely, in 23% of the individual studies, the control group outperformed the experimental group as measured by the differences in adjusted assessment scores between the two groups (Marzano & Haystead, 2010).

Although the first phase findings suggested, on average, a positive effect on student achievement, the percentage of individual studies with a negative effect size indicated that those teachers were less effective in their implementation of the interactive technologies. Videotapes were analyzed during the second phase of the evaluation study to determine how well teachers demonstrated specific instructional strategies while teaching the experimental group of students.

In the second phase of analysis, a strong relationship was found between the magnitude and direction of effect sizes that were calculated in the first phase and teachers' implementation of six instructional strategies (chunking of content, scaffolding of content, pacing of instruction, monitoring of student progress, clarity of content, and increases in student response rates), $R(98) = .79$, $p < .001$. It is worth noting that 60% of the variability in the effect sizes was accounted for by these six variables, $R^2 = .62$ (Marzano & Haystead, 2010).

A reasonable inference can be made that interactive technology would be expected to have a greater influence on student achievement when teachers improve their fidelity in their instructional practices. As evidenced by the second phase findings reported by Marzano and Haystead (2010), students of teachers who combined effective instructional practices with interactive technology outperformed students in the control group to a greater extent than the students of teachers who merely utilized the technology in their lessons.

This relationship between effective instructional practices and the use of interactive technology to improve student achievement inspired action research at an elementary school in Southern California.

Method

In July 2009, the school's principal and school district's director of instructional technology met with the second author and expressed an interest in examining the extent to which professional development focused on using technology to enhance the *Art and Science of Teaching* framework (Marzano, 2007) might improve student achievement and engagement. Action research began at the school during the 2009–2010 academic year and continued through the 2011–2012 academic year.

As mentioned previously, this report presents a descriptive case study of the action research through visual analyses of school proficiency data from statewide student assessments from 2004 through 2012. In addition, this report presents statistical analyses of school and district proficiency data from 2009 and 2012. Again, the descriptive case study sought to answer the following research question:

> Does teacher professional development focused on using interactive technologies to enhance instructional strategies (articulated in the *Art and Science of Teaching* framework; Marzano, 2007) influence student engagement and achievement?

Every classroom in the school had IWBs and clickers. It should be noted that the IWBs and clickers were new to the teachers at the start of their action research. Therefore, a key consideration was how to simultaneously build teachers' capacity with the instructional strategies and the interactive technology.

Figure 2 shows the theory of action for the descriptive case study.

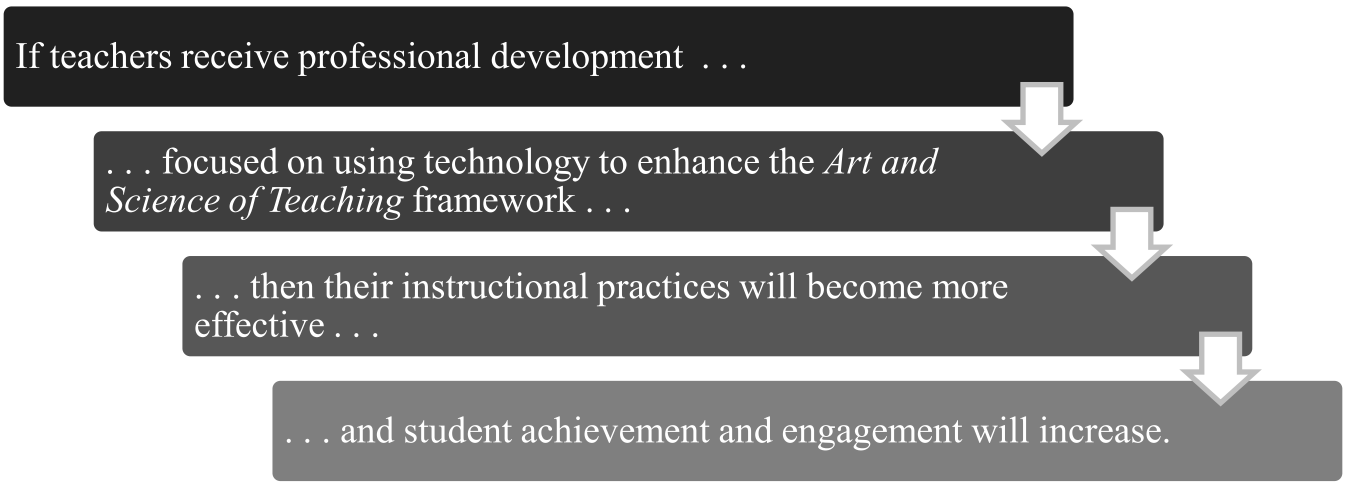

Figure 2: Theory of action for the descriptive case study.

Previous evaluation studies informed the theory of action for the descriptive case study. It seemed a reasonable supposition that if teachers received professional development focused on using interactive technology to enhance the instructional strategies in the *Art and Science of Teaching* framework (Marzano, 2007), then teachers' instructional effectiveness would increase, which would lead to increases in student engagement and achievement.

The school's and district's reports of state and federal accountability were accessed through the California Department of Education's (CDE) *DataQuest* website (2013a). One metric used for state accountability is the Academic Performance Index (API) score. The API score is calculated by converting a student's performance on assessments from California's Standardized Testing and Reporting (STAR) program across multiple content areas into points on the API scale. Assessments are administered to students in grades 2 through 11.

An API score is calculated for schools, districts, and student groups with 11 or more valid assessment scores, with points averaged across all students and tests. To measure academic improvement, each reporting cycle consists of a Growth API score which is calculated from the current year's assessments and a Base API score which is calculated using results from the previous year. It is important to note that the Base API score in a reporting cycle might reflect an adjustment to the Growth API score that was reported the previous year. In other words, due to changes in assessments and weights used to calculate the Growth API score in a specific reporting cycle, the Base API score might be higher or lower than the Growth API score that was included in the previous year's accountability report. Therefore, direct comparisons of either Base API or Growth API scores across reporting cycles are considered by CDE to be invalid. However, the amount of growth (change) in a reporting cycle can be compared across reporting cycles (Growth API score – Base API score). In addition, API scores can be meaningfully compared across subgroups within a single reporting cycle (CDE, 2013b).

Student academic achievement was measured at the school level by API scores and the percentages of students, grades 2 through 6, who scored at or above the state's designated level of proficiency on STAR English language arts and mathematics assessments (hereinafter referred to as *proficient*). Although academic achievement data was limited to statewide student assessments for grades 2 through 6, the action research involved all students and teachers at the school, including a self-contained classroom for special-needs students with a variety of special learning needs ranging from moderate to severe.

As mentioned earlier, the school employed 22 teachers and served 513 students. The students' demographics were reasonably representative of the larger community population (60% Latino, 19% African American, 11% Caucasian, 2% Asian, 1% Pacific Islander, and 1% Filipino). In the school, 85% of the students qualified for Title 1 services (such as free or reduced lunch), 82% of the students were racial or ethnic minority learners (mostly Latino and African American), and 51% of the students were English learners (ELs). Additionally, 24% of the students were considered *highly mobile*, meaning they moved their place of residence at least twice during an academic year.

The professional development team began implementing the intervention with the school's teachers in January 2010 and continued to work with the instructional staff for the next three years. The professional development focused on how to use IWBs and clickers to support and enhance the strategies articulated in the *Art and Science of Teaching* framework (Marzano, 2007). Time was given for teachers to practice, deepen, elaborate, and revise their knowledge and skills, as well as to reflect on their learning. Individual and collective gains were recognized and celebrated.

The professional development team met with the teachers once a month. This progression continued for the first year and then tapered off during the second and third years as teachers' confidence and capacity developed. Each professional development session in the progression was designed to increase teachers' understanding and application of the critical design questions in the *Art and Science of Teaching* framework (Marzano, 2007). In addition, the sessions were designed to show teachers how to use their existing technology to improve their implementation of specific strategies.

It should be noted that teachers were separated according to their comfort level with technology. Teachers who self-assessed as *beginning* or *applying* were in one group and teachers who self-assessed as *emerging* or *innovating* were in a second group. This grouping seemed to increase teachers' comfort level and allow them to learn and practice at their own pace.

Results

Student academic achievement was measured at the school level with API scores and the percentages of students, grades 2 through 6, who scored *proficient* on STAR English language arts (ELA) and mathematics (math) assessments.

Figure 3 displays a chart of the school's Base and Growth API scores for each reporting cycle from 2004 to 2012.

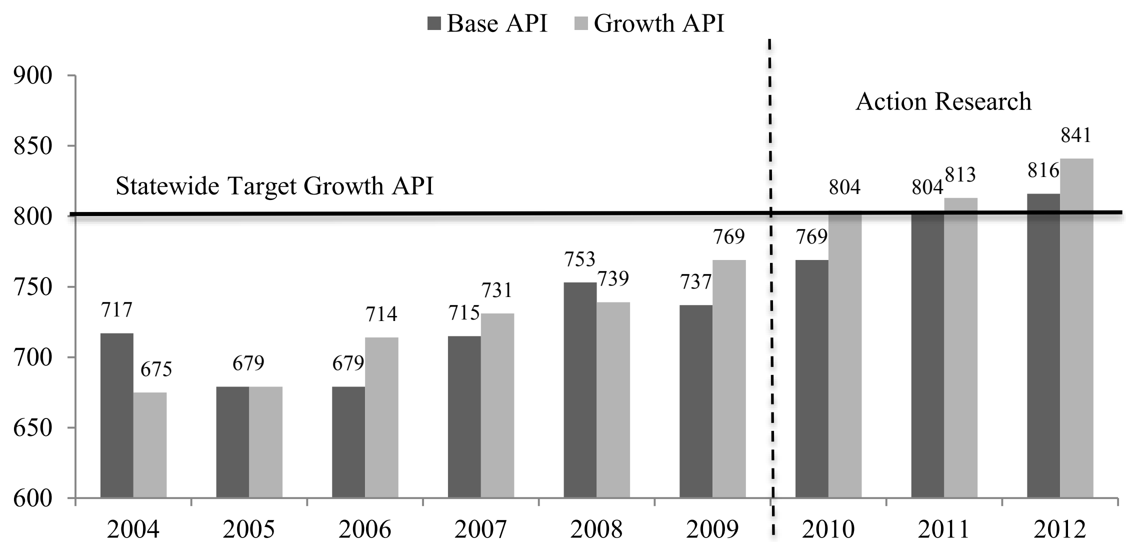

Figure 3: Base and Growth API scores at the school from 2004 to 2012.

Source. California Department of Education, 2013a.

As Figure 3 shows, the school reached the statewide target Growth API score of 800 in 2010, the first year the professional development was implemented.

For reference purposes, the school's annual accountability report also includes median values for Base and Growth API scores from 100 schools with similar demographic characteristics (for example, pupil ethnicity, pupil socioeconomic status, percentage of pupils who are English learners, and so on). Figure 4 shows the school and median Growth API scores in each reporting cycle from 2004 to 2012.

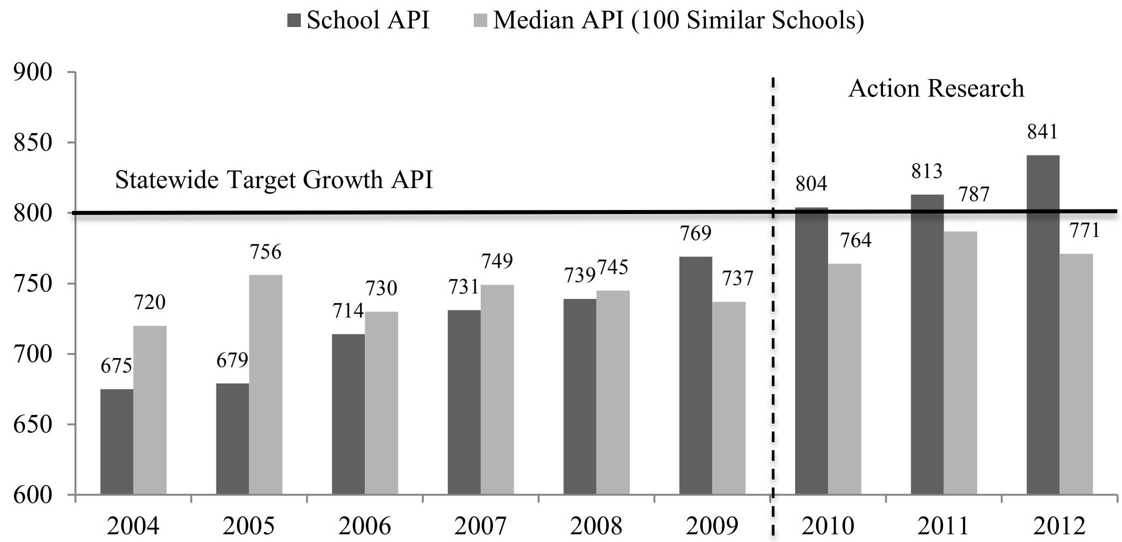

Figure 4: A comparison of school and median Growth API scores from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 4 indicates that the school's Growth API score was lower than the median score calculated from 100 similar schools in the reporting cycles from 2004 to 2008. In 2009, one year before teachers started receiving the focused professional development, the school's Growth API score was higher than the median score, a positive trend that continued throughout the three-year period of action research.

As mentioned previously, one of the valid ways to directly compare API scores from one reporting cycle to another is the amount of change between Base and Growth API scores in each reporting cycle. Figure 5 displays the changes in Base and Growth API scores for each reporting cycle from 2004 to 2012.

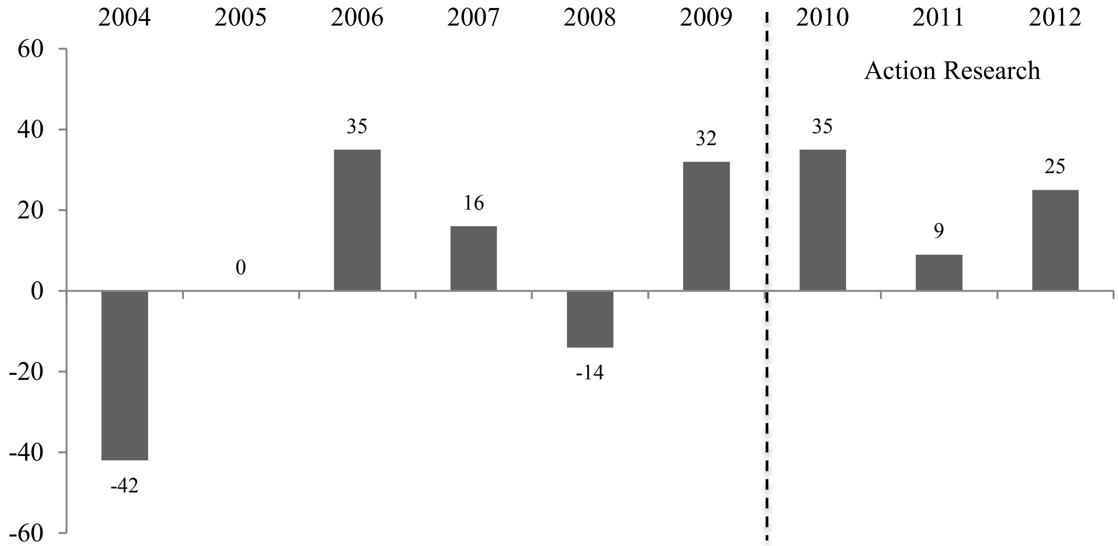

Figure 5: Changes in Base and Growth API scores at the school from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 5 indicates that in the six years before the school's action research (2004–2009), the amount of change in Base and Growth API scores was negative in two reporting cycles and positive in three reporting cycles. During the period of action research (2010–2012), the school maintained a positive change in Base and Growth API scores in each reporting cycle. Means (with standard deviations in parentheses) for each period were 7.00 (27.71) and 23.00 (13.11), respectively.

Figure 6 displays a comparison of the school's change in Base and Growth API scores and the median change in Base and Growth API scores for 100 similar schools in each reporting cycle from 2004 to 2012.

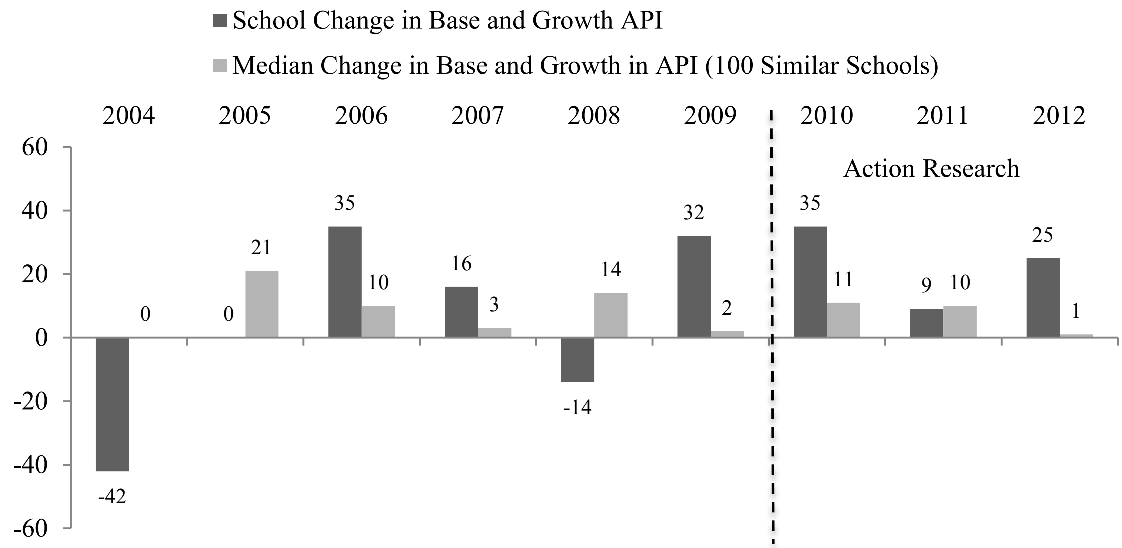

Figure 6: A comparison of school and median changes in Base and Growth API scores from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 6 illustrates that in the six years before the action research (2004–2009), the school's change in Base and Growth API scores was lower than the median calculated from 100 similar schools in three of the six reporting cycles and higher than the median in three of the cycles. During the period of action research (2010–2012), the school's change in Base and Growth API scores was higher than the median in two of the three reporting cycles. It is worth mentioning that in 2011, the school's change in Base and Growth API scores was only one point lower than the median. Means (with standard deviations in parentheses) calculated from the median values for each period were 13.43 (15.40) and 7.33 (5.51), respectively. Although not statistically significant at the standard alpha of .05 (that is, $p < .05$), the school's means for each period were 6.43 points lower and 15.67 points higher, respectively, $t(12) = –0.54$, $p = .601$ and $t(4) = 1.91$, $p = .129$.

Again, API scores are calculated from students' scores on STAR assessments across multiple content areas. Therefore, Figure 7 displays the percentages of students who scored *proficient* on the STAR ELA and math assessments from 2004 to 2012.

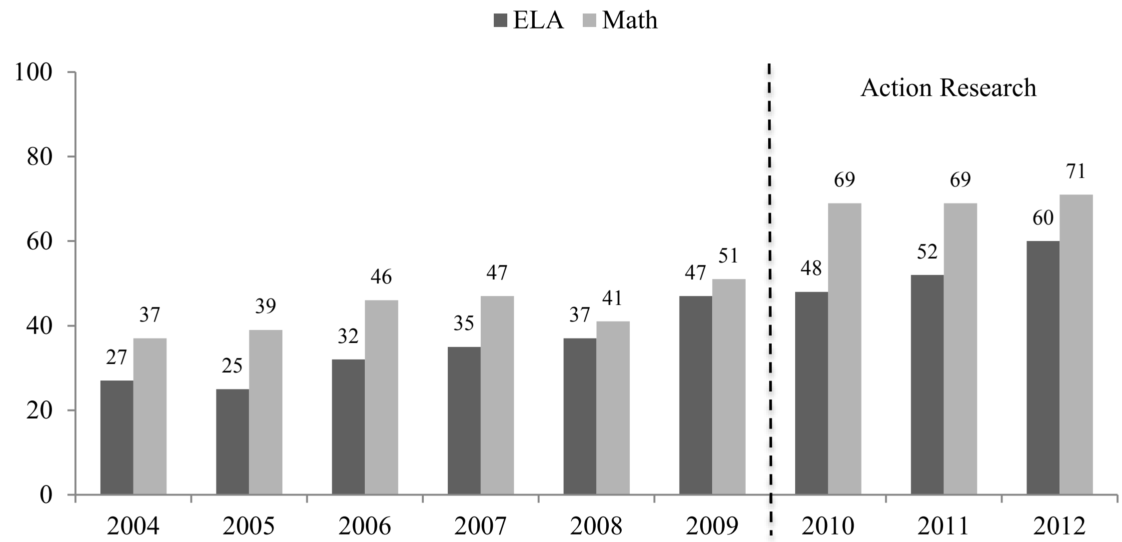

Figure 7: Percentages of students at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 7 indicates a positive trend in both ELA and math from 2004 to 2012. It is important to note that during the first year of action research (2010), the percentage of students who scored *proficient* on the STAR math assessment increased eighteen points from 51% (2009) to 69%. The percentages were close to 70% for the entire three-year period of action research. It is also worth noting that, although there was little change between 2009 and 2010, during the period of action research the percentage reported for STAR ELA increased twelve points from 48% in 2010 to 60% in 2012. Therefore, a comparison was also made between categories of student groups (for example, ethnicity, socioeconomic status, English language acquisition, and so on).

Figures 8 through 10 display the percentages of African American, Latino, and Caucasian students who scored *proficient* on the STAR ELA and math assessments from 2004 to 2012.

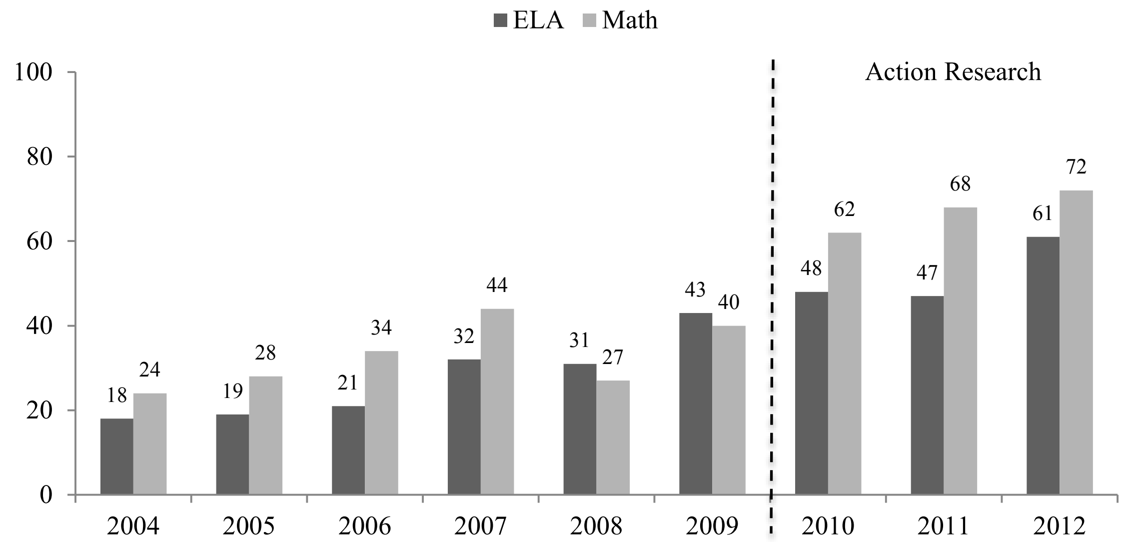

Figure 8: Percentages of African American students at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

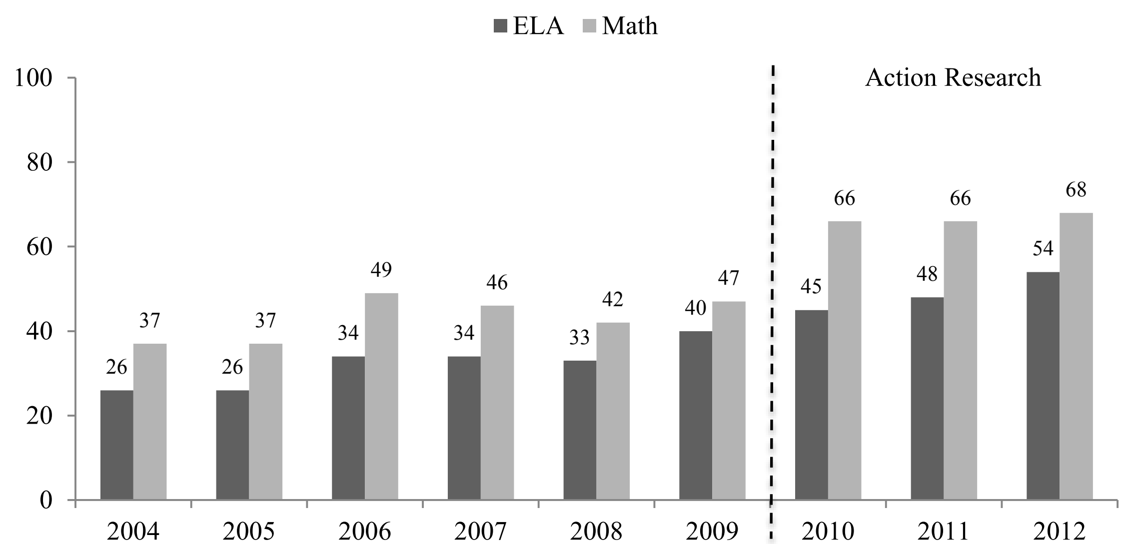

Figure 9: Percentages of Latino students at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

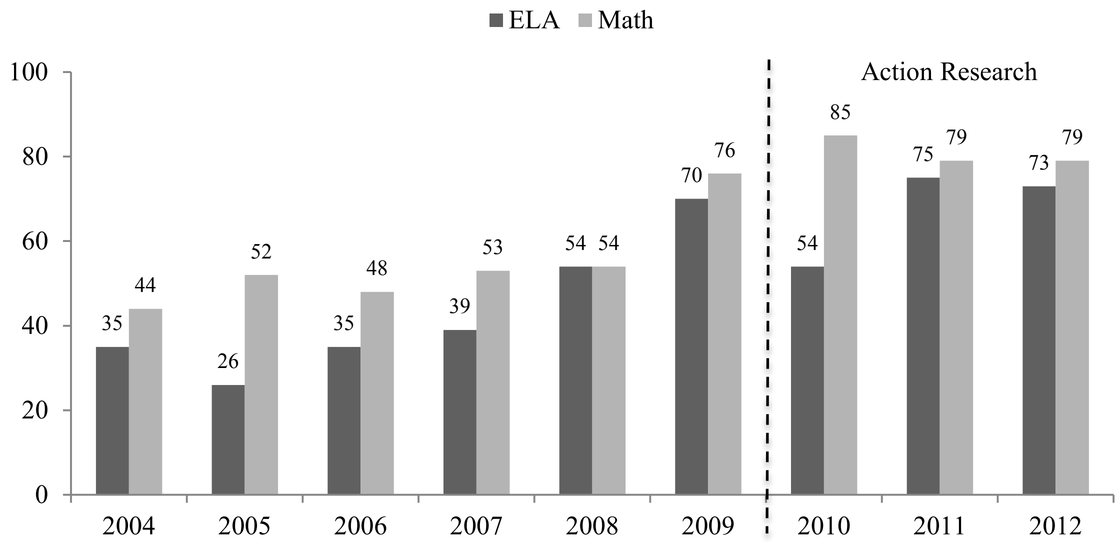

Figure 10: Percentages of Caucasian students at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 8 (page 12) shows that, in the first year of action research, the percentage of African American students who scored *proficient* on the STAR ELA assessment increased five points from 43% in 2009 to 48% in 2010. In the final year of action research (2012), the percentage had increased thirteen points to 61%. In the first year of action research (2010), the percentage reported for STAR math increased twenty-two points from 40% (2009) to 62%. By the end of the period of action research (2012), the percentage increased ten points to 72%.

Figure 9 (page 12) illustrates similar increases in percentages of Latino students who scored *proficient* on the STAR ELA and math assessments. In 2010, the percentage reported for STAR ELA increased five points from 40% (2009) to 45% and the percentage reported for STAR math increased nineteen points from 47% (2009) to 66%. In 2012, the percentage reported for STAR ELA had increased nine more points to 54% and the percentage reported for STAR math remained about the same with a two-point increase to 68%.

Finally, Figure 10 indicates that prior to the school's period of action research, the percentage of Caucasian students who scored *proficient* on the STAR ELA assessment increased sixteen points from 54% in 2008 to 70% in 2009. A larger increase was reported for STAR math in the same year, twenty-two points from 54% to 76%. In the first year of action research (2010), the percentage reported for STAR ELA decreased back to 54%. However, in 2011 the percentage increased twenty-one points to 75% and remained about the same in 2012, a two-point decrease to 73%. In comparison, the percentage reported for STAR math increased nine points to 85% in 2010 and decreased six points to 79% in both 2011 and 2012.

Figures 11 through 13 show the percentages of low-income students, English learners, and students with disabilities who scored *proficient* on the STAR ELA and math assessments from 2004 to 2012.

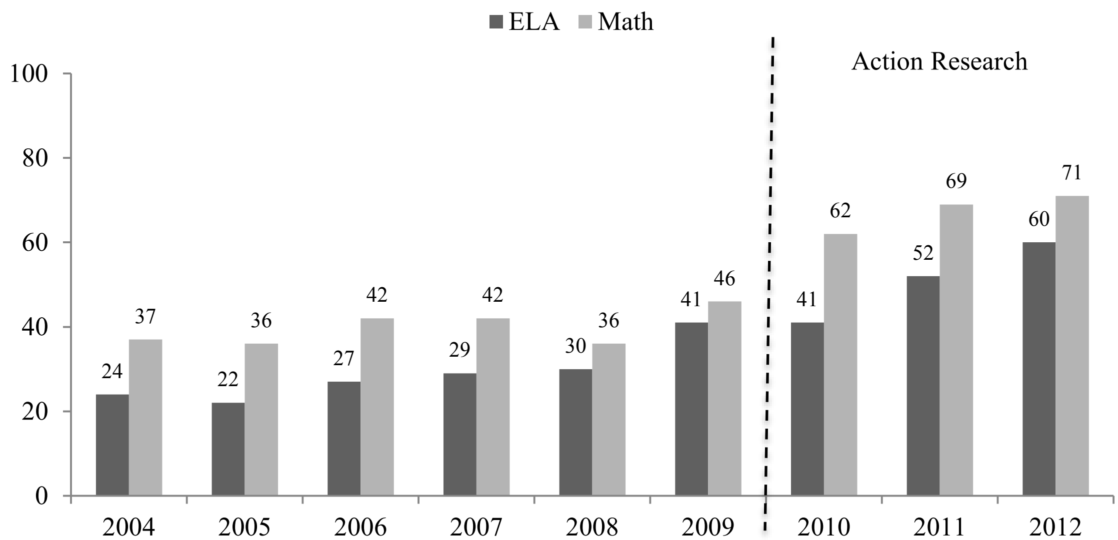

Figure 11: Percentages of low-income students at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

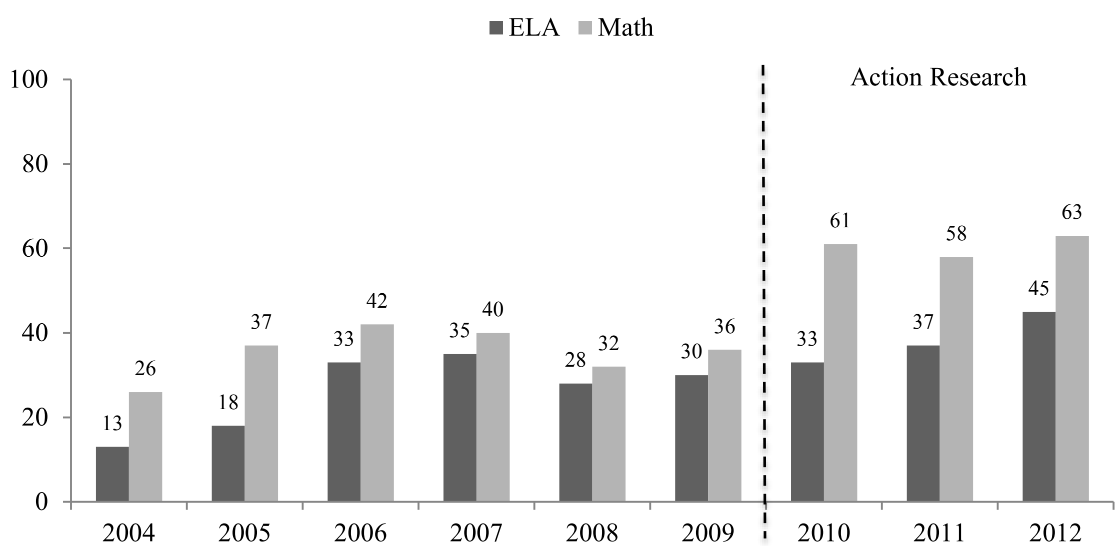

Figure 12: Percentages of English learners at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

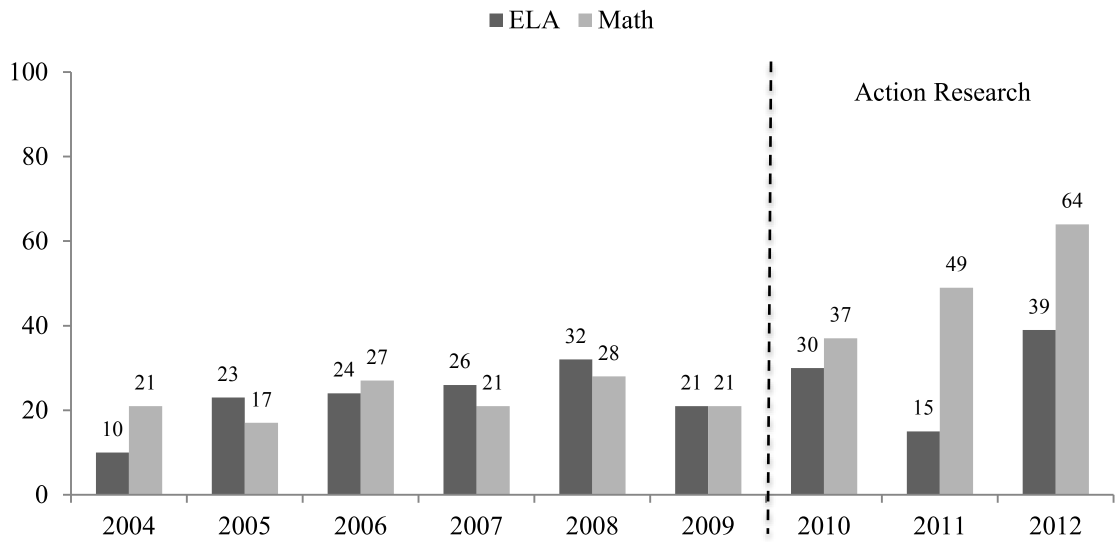

Figure 13: Percentages of students with disabilities at the school who scored proficient on STAR ELA and math assessments from 2004 to 2012.

Source. California Department of Education, 2013a.

Figure 11 (page 14) shows that, in the first year of action research (2010), the percentage of low-income students who scored *proficient* on the STAR ELA assessment was 41%, the same as in the previous year. By the end of the action research period (2012), the percentage increased nineteen points to 60%. In comparison, the percentage reported for STAR math increased sixteen points from 46% in 2009 to 62% in 2010. In 2012, the percentage had increased nine points to 71%.

Figure 12 (page 14) illustrates that, in the first year of action research (2010), the percentage of English learners who scored *proficient* on the STAR ELA assessment increased three points from 30% in 2009 to 33%. In the final year of action research (2012), the percentage increased twelve points to 45%. In comparison, the percentage reported for STAR math increased twenty-five points from 36% in 2009 to 61% in 2010. During the period of action research, the percentages remained close to 60% with 58% in 2011 and 63% in 2012.

Finally, Figure 13 indicates that, in the first year of action research (2010), the percentage of students with disabilities who scored *proficient* on the STAR ELA assessment increased nine points from 21% in 2009 to 30%. In 2011, the percentage decreased fifteen points to 15%. In 2012, the percentage increased twenty-four points to 39%. In comparison, the percentage reported for STAR math increased sixteen points from 21% in 2009 to 37% in 2010. The percentage continued to grow during the period of action research, increasing twelve points to 49% in 2011 and another fifteen points to 64% in 2012.

Figures 7 through 13 presented a visual analysis of the school's percentages of students who scored *proficient* on the STAR ELA and math assessments. In order to conduct a statistical analysis of the school's proficiency data with a control reference point, CDE's website (2013a) was used to obtain proficiency data from five elementary schools located in the school's district (labeled District 1 through District 5). The statistical software package IBM SPSS Statistics was used to obtain results from Pearson's chi-square test comparing school and district proficiency data in 2012, the final year of the school's three-year period of action research. School and district proficiency data were also compared in 2009, the year before the school's action research, to provide a baseline reference prior to the school's implementation of focused professional development. Pearson's chi-square provides a mechanism to statistically compare the frequencies that are observed for a specific category with the frequencies that might be expected by chance (see Technical Note 1).

Table 1 displays the number of students who scored *proficient* and *below proficient* in 2009.

Table 1: School and District Proficiency Counts for STAR ELA and Math in 2009

	STAR ELA		STAR Math	
	Proficient	**Below Proficient**	**Proficient**	**Below Proficient**
School	119 (46.7%)	136 (53.3%)	130 (51.0%)	125 (49.0%)
District 1	143 (53.4%)	125 (46.6%)	147 (54.9%)	121 (49.3%)
District 2	108 (43.2%)	142 (56.8%)	104 (41.8%)	145 (58.2%)
District 3	93 (47.9%)	101 (52.1%)	114 (58.8%)	80 (41.2%)
District 4	143 (46.1%)	167 (53.9%)	160 (51.6%)	150 (48.4%)
District 5	171 (51.2%)	163 (48.8%)	162 (48.5%)	172 (51.5%)

Table 1 indicates that, in the year prior to the school's period of action research, the proportion of students in the school who scored *proficient* on the STAR ELA assessment was lower than three of the schools in the district. The proportion of students in the school who scored *proficient* on the STAR math assessment was also lower than three schools in the district.

Table 2 lists the results of the chi-square tests.

Table 2: Pearson's Chi-Square Test Results for STAR ELA and Math in 2009

	STAR ELA			STAR Math		
Comparison	c^2	N	p	c^2	N	p
School vs. District 1	2.34	523	.126	0.79	523	.375
School vs. District 2	0.61	505	.434	4.30*	504	.038
School vs. District 3	0.71	449	.789	2.69	449	.101

Comparison	STAR ELA			STAR Math		
	c^2	*N*	*p*	c^2	*N*	*p*
School vs. District 4	0.16	565	.899	0.02	565	.881
School vs. District 5	1.19	589	.276	0.36	589	.551

$^*p < .05$.

Table 2 shows that only one comparison (School vs. District 2, STAR Math) would be considered statistically significant at the standard $\alpha = .05$, $\chi^2(1) = 4.30$, $N = 504$, $p = .038$. At $\alpha = .05$ (that is, $p < .05$), assuming the null hypothesis of no association between the two categorical variables to be true, the conditional probability of mistakenly rejecting the null hypothesis would be less than 5% (see Technical Note 2). Taken at face value, with the one exception, these findings suggest that the school's proficiency rates for the STAR ELA and math assessments in 2009 were very similar to the other elementary schools in the district.

Table 3 displays the number of students who scored *proficient* and *below proficient* in 2012.

Table 3: School and District Proficiency Counts for STAR ELA and Math in 2012

	STAR ELA		STAR Math	
	Proficient	Below Proficient	Proficient	Below Proficient
School	200 (59.7%)	135 (40.3%)	238 (71.0%)	97 (29.0%)
District 1	198 (57.6%)	146 (42.4%)	223 (64.8%)	121 (35.2%)
District 2	158 (46.7%)	180 (53.3%)	180 (53.3%)	158 (46.7%)
District 3	143 (57.9%)	104 (42.1%)	161 (65.2%)	86 (34.8%)
District 4	203 (52.3%)	185 (47.7%)	246 (63.4%)	142 (36.6%)
District 5	224 (50.5%)	220 (49.5%)	227 (51.2%)	216 (48.8%)

Table 3 indicates that the proportions of students in the school who scored *proficient* on the STAR ELA and math assessments in the final year of the school's action research were higher than the five elementary schools in the district.

Table 4 reports the results of the chi-square tests.

Table 4: Pearson's Chi-Square Test Results for STAR ELA and Math in 2012

Comparison	STAR ELA			STAR Math		
	c^2	*N*	*p*	c^2	*N*	*p*
School vs. District 1	0.32	679	.571	3.01†	679	.083
School vs. District 2	11.34***	673	<.001	22.63***	673	<.001
School vs. District 3	0.19	582	.661	2.27	582	.132
School vs. District 4	3.97*	723	.046	4.75*	723	.029
School vs. District 5	6.59*	779	.010	31.11***	778	<.001

†$p < .10$. *$p < .05$. ***$p < .001$.

Table 4 shows that six comparisons would be considered statistically significant at the standard α = .05:

- STAR ELA
  - School vs. District 2, $\chi^2(1) = 11.34$, $N = 673$, $p < .001$
  - School vs. District 4, $\chi^2(1) = 3.97$, $N = 723$, $p = .046$
  - School vs. District 5, $\chi^2(1) = 6.59$, $N = 779$, $p = .010$
- STAR Math
  - School vs. District 2, $\chi^2(1) = 22.63$, $N = 673$, $p < .001$
  - School vs. District 4, $\chi^2(1) = 4.75$, $N = 723$, $p = .029$
  - School vs. District 5, $\chi^2(1) = 31.11$, $N = 778$, $p < .001$

In addition, one comparison (School vs. District 1, STAR Math) would be considered significant at the less conservative α = .10, $\chi^2(1) = 3.01$, $N = 679$, $p = .083$. Taken at face value, these findings suggest that in three out of five comparisons, there was a significant association between the implementation of focused professional development and whether or not students scored *proficient* on the STAR ELA assessment. The findings also suggest that in four out of five comparisons, there was a significant association between the implementation of focused professional development and whether or not students scored *proficient* on the STAR math assessment.

In order to examine the magnitude of association between the focused professional development and student proficiency rates, odds ratio and rate ratio (or relative risk) effect sizes were also calculated (see Technical Note 3). Tables 5 and 6 display the odds ratio and rate ratio effect sizes for STAR ELA and math in 2012 along with their respective 95% confidence intervals. The 95% confidence interval represents the precision of the calculated effect size and indicates a 95% probability that the true estimated effect size is within the lower and upper limits of the interval. When the interval does not include 1.00, the odds ratio and rate ratio effect sizes would be considered statistically significant at the standard α = .05.

Table 5: Odds Ratio Effect Sizes for STAR ELA and Math in 2012

	STAR ELA		STAR Math	
Comparison	*OR*	95% CI	*OR*	95% CI
School vs. District 1	1.09	[0.81, 1.48]	1.33	[0.96, 1.84]
School vs. District 2	1.69	[1.24, 2.29]	2.15	[1.57, 2.96]
School vs. District 3	1.08	[0.77, 1.50]	1.31	[0.92, 1.87]
School vs. District 4	1.35	[1.01, 1.81]	1.42	[1.04, 1.94]
School vs. District 5	1.46	[1.09, 1.94]	2.34	[1.73, 3.15]

Note. *OR* = odds ratio effect size; CI = confidence interval.

Table 5 indicates that all of the odds ratio effect sizes were greater than 1.00. Taken at face value, the effect sizes suggest that the odds of students at the school scoring *proficient* on the STAR ELA and math assessments were higher than students at the other elementary schools in the district. The statistical software package Comprehensive Meta-Analysis was used to calculate weighted mean estimates of the odds ratio effect size for STAR ELA and math using a random-effects model of error (Lipsey & Wilson, 2001). Meta-analytic techniques are often used to combine results of statistical analyses in order to calculate a more reliable estimate of the effect of a specific intervention. The findings were statistically significant at $\alpha = .001$:

- STAR ELA: *OR* = 1.32, $p < .001$, 95% CI [1.12, 1.56]
- STAR Math: *OR* = 1.67, $p < .001$, 95% CI [1.30, 2.14]

The mean estimates of effect size indicate that the odds of students at the school scoring *proficient* on the STAR ELA and math assessments were 1.32 and 1.67 times higher, respectively, than students at the other elementary schools in the district.

Table 6: Rate Ratio Effect Sizes for STAR ELA and Math in 2012

	STAR ELA		STAR Math	
Comparison	*RR*	95% CI	*RR*	95% CI
School vs. District 1	1.04	[0.91, 1.18]	1.10	[0.99, 1.22]
School vs. District 2	1.28	[1.11, 1.48]	1.33	[1.18, 1.51]
School vs. District 3	1.03	[0.90, 1.18]	1.09	[0.97, 1.22]
School vs. District 4	1.14	[1.00, 1.30]	1.12	[1.01, 1.24]
School vs. District 5	1.18	[1.04, 1.34]	1.39	[1.24, 1.55]

Note. *RR* = rate ratio effect size; CI = confidence interval.

Table 6 shows that all of the rate ratio effect sizes were greater than 1.00. Taken at face value, the effect sizes suggest that the probabilities of students at the school scoring *proficient* on the STAR ELA and math assessments were greater than students at the other elementary schools in the district. Weighted mean estimates of the rate ratio effect size for STAR ELA and math were also calculated. The findings were statistically significant at $\alpha = .01$ and $\alpha = .001$:

- STAR ELA: $RR = 1.13$, $p = .002$, 95% CI [1.04, 1.22]
- STAR Math: $RR = 1.20$, $p < .001$, 95% CI [1.09, 1.32]

The mean estimates of effect size indicate that students at the school were 13% and 20% more likely to score *proficient* on the STAR ELA and math assessments, respectively, than students at the other elementary schools in the district.

Summary and Discussion

Teachers in a Southern California, high-needs, low-achieving elementary school participated in a three-year period of action research beginning in the 2009–2010 academic year. As part of the action research, teachers received 40 hours of in-person professional development, classroom coaching, and online support to improve their understanding and application of interactive technologies (that is, IWBs and clickers) to enhance their utilization of instructional strategies articulated in the *Art and Science of Teaching* framework (Marzano, 2007).

This report presented a descriptive case study of the action research through visual analyses of school proficiency data from statewide student assessments from 2004 through 2012. In addition, this report presented statistical analyses of school and district proficiency data from 2009 and 2012.

The descriptive case study sought to address the following research question:

> Does teacher professional development focused on using interactive technologies to enhance instructional strategies (articulated in the *Art and Science of Teaching* framework; Marzano, 2007) influence student engagement and achievement?

The visual analyses of the school's Academic Performance Index (API) scores revealed that the school reached the statewide target Growth API score of 800 in 2010, during the first year of the action research. In addition, the school's Growth API score was higher than the median value calculated from 100 schools with similar demographic characteristics (such as pupil ethnicity, pupil socioeconomic status, and so on). In the six-year period before the school's action research, the school's mean change in Base and Growth API scores (with standard deviation in parentheses) was 7.00 (27.71) compared with a mean change of 23.00 (13.11) during the three-year period of action research. Taken at face value, the difference between the mean change in Base and Growth API scores in each period suggests that the focused professional development had an influence on student achievement.

The visual analyses of the school's student proficiency counts indicated an eighteen point increase in the percentage of students who scored *proficient* on the STAR math assessment during the first year of action research from 51% in 2009 to 69% in 2010. Although there was little change in the percentage of students who scored *proficient* on the STAR ELA assessment between 47% in 2009 and 48% in 2010, the reported percentage increased twelve points to 60% in 2012, the final year of action research.

A comparison of percentages for African American, Latino, and Caucasian students revealed that there was a sixteen-point increase in the percentage of Caucasian students who scored *proficient* on the STAR math assessment from 54% in 2008 to 76% in 2009, one year prior to the school's period of action research. In the first year of action research (2010), the percentage reported for STAR math increased nine points to 85%. By the end of the action research (2012), the percentage decreased six points to 79%. In comparison, the percentage of African American students who scored *proficient* on the STAR math assessment increased twenty-two points from 40% in 2009 to 62% in 2010. By the end of the action research (2012), the percentage increased ten points to 72%. The percentage of Latino students who scored *proficient* on the STAR math assessment increased nineteen points from 47% in 2009 to 66% in 2010. In 2012, the percentage increased two points to 68%. Smaller percentage increases were reported for the STAR ELA assessment. Taken at face value, the comparison between ethnic categories suggests that the focused professional development had a greater influence on African American and Latino students.

A comparison of percentages was also made for low-income students, English learners, and students with disabilities. It is worth noting that all three categories of students exhibited fluctuating percentages that were below 50% in the six years before the school's action research. During the entire period of action research, the reported percentages for low-income students and English learners were above 50% for the STAR math assessment. Although students with disabilities only had a percentage above 50% during the final year of action research, it is important to consider that the percentage increased twelve points from 37% in 2010 to 49% in 2011 and fifteen points to 64% in 2012, a total increase of almost thirty points. It is also worth considering that the percentages of low-income students, English learners, and students with disabilities who scored *proficient* on the STAR math assessment increased 25%, 27%, and 43%, respectively from 2009 to 2012. Smaller increases were reported for the STAR ELA assessment. Taken at face value, the comparison between these categories of student classification suggests that the focused professional development had a considerable influence on math proficiency percentages in all three categories, with the largest increase reported for students with disabilities.

Again, CDE's website (2013a) was used to obtain proficiency data from five elementary schools located in the school's district (labeled District 1 through District 5) to provide a control reference point for statistical analyses of the school's proficiency data. Pearson's chi-square test indicated that, at the end of the three-year period of action research, six comparisons would be considered statistically significant at the standard $\alpha = .05$ (that is, $p < .05$):

- STAR ELA
  - School vs. District 2, $\chi^2(1) = 11.34$, $N = 673$, $p < .001$
  - School vs. District 4, $\chi^2(1) = 3.97$, $N = 723$, $p = .046$
  - School vs. District 5, $\chi^2(1) = 6.59$, $N = 779$, $p = .010$
- STAR Math
  - School vs. District 2, $\chi^2(1) = 22.63$, $N = 673$, $p < .001$
  - School vs. District 4, $\chi^2(1) = 4.75$, $N = 723$, $p = .029$
  - School vs. District 5, $\chi^2(1) = 31.11$, $N = 778$, $p < .001$

In addition, one comparison (School vs. District 1, STAR Math) would be considered significant at the less conservative $\alpha = .10$, $\chi^2(1) = 3.01$, $N = 679$, $p = .083$.

Taken at face value, the results obtained from Pearson's chi-square tests suggest that in three out of five comparisons for STAR ELA, and in four out of five comparisons for STAR math, there was a significant association between the implementation of focused professional development to enhance teachers' utilization of instructional strategies articulated in the *Art and Science of Teaching* framework (Marzano, 2007) and whether or not students scored *proficient* on the assessment.

These findings are further supported by the weighted mean estimates of the odds ratio and rate ratio effect sizes for STAR ELA and math:

- STAR ELA
  - $OR = 1.32$, $p < .001$, 95% CI [1.12, 1.56]
  - $RR = 1.13$, $p = .002$, 95% CI [1.04, 1.22]
- STAR Math
  - $OR = 1.67$, $p < .001$, 95% CI [1.30, 2.14]
  - $RR = 1.20$, $p < .001$, 95% CI [1.09, 1.32]

The odds ratio effect size for STAR ELA indicates that the odds of scoring *proficient* on the assessment were 1.32 times higher if students were taught by teachers who participated in the focused professional development than if taught by teachers who did not receive the professional development. The rate ratio of 1.13 suggests that students were 13% more likely to score *proficient* on the STAR ELA assessment if taught by teachers who received the professional development. The odds ratio effect size of 1.67 for STAR math suggests that the odds of scoring *proficient* on the assessment were 1.67 times higher if students were taught by teachers who received the professional development than if taught by teachers who did not receive the professional development. The rate ratio effect size of 1.20 suggests that students were 20% more likely to score *proficient* on the STAR math assessment if taught by teachers who received the professional development.

It is worth noting that in the year prior to the school's action research, there was only one comparison (School vs. District 2, STAR Math) that would be considered statistically significant at the standard $\alpha = .05$, $\chi^2(1) = 4.30$, $N = 504$, $p = .038$. By the end of the period of action research, the odds ratio effect size increased from *OR* = 1.45, $p = .038$, 95% CI [1.02, 2.06], to *OR* = 2.15, $p < .001$, 95% CI [1.57, 2.96]. The rate ratio effect size increased to a lesser degree from *RR* = 1.22, $p = .038$, 95% CI [1.01, 1.48], to *RR* = 1.33, $p < .001$, 95% CI [1.18, 1.51].

Given that, with the one exception (School vs. District 2, STAR Math), there were no significant associations between proficiency counts at the school and the other elementary schools in the district in the year prior to the action research, a reasonable inference can be made that the focused professional development was significantly related to the difference between school and district proficiency counts for STAR ELA and math.

In September 2011, a video-taped interview was conducted with the school's principal and district director of instructional technology. At one point in the interview, the school principal reflected on his concerns with student attendance, student academic performance, and student discipline prior to the action research:

> Like many urban schools, our school was struggling with low student attendance, low student academic performance, and high discipline problems. As a matter of fact, we were at the bottom of our district in terms of student achievement and attendance. Teacher engagement and teacher satisfaction was also incredibly low. The future did not look very bright because the trajectory was heading in the wrong direction.

When asked whether student disciplinary issues were impacted by the pedagogical and technological changes evident in teacher instruction, the school principal responded:

> Student discipline issues dropped significantly during the three years of this study. While we don't keep records of student discipline, as the building principal, I deal with these issues personally, and I have noticed a significant drop in the number of students who are referred to me for behavior issues. I think this is because students are more engaged in their lessons and as a result they are less likely to become frustrated and act out.

The school principal also reflected on the influence of the action research on student achievement, student engagement, and student attendance:

> After a couple of years into this initiative, the achievement gap among [racial and ethnic] minority students has definitely narrowed. I think this has a lot to do with engagement. The engagement in our classrooms is now off the charts. Our kids are engaged [and] our teachers are engaged in what they're teaching. They're enthusiastic and excited and our gains in student achievement are the result. In 2009, we had the worst attendance rate in the district and now we're actually the school with the highest attendance in the district. I see the engagement part not only impacting achievement, but also impacting attendance. We're very pleased with the results that we're getting. We still have a lot of work to do, but we're heading in the right direction.

When considering the achievement gap, the school principal elaborated:

> There is a lot of conversation about the achievement gap. It's an issue that not only our schools, but our society, is facing: students of color and students whose English language is at a developmental level predictably score lower than their White and Asian counterparts. What we've really been looking at here at our school is that the achievement gap is really the result of an engagement gap. . . . All students come to school wanting to learn and wanting to be engaged. So that's a given. All students come with that set of goals. So when we really look at the engagement gap, we're actually focusing more on a teaching gap. By this, what I'm saying is that when teachers have the skills, they have the resources, and they have the professional development to engage all students, then we're able to see the achievement gap get addressed, which is a completely different focus than what we generally do in public education where we focus on the child and their [sic] deficits. Now we're actually shifting the focus to the teacher, providing them [sic] with the skills and the resources and the support, all the way from professional development to school-wide support, to actually engage all kids.

When asked about teacher attitudes toward the action research, the school principal answered:

> The teacher response was overwhelmingly positive—they couldn't get enough of it. They immediately saw the potential of the technology, but knew that they needed to learn how to maximize that technology. The professional development was engaging and exciting and it took teachers from where they were to the next step. Because this became a part of the school culture, everyone was moving forward. This project has brought our staff together and created more collegial relationships between all of our teachers. It was actually only a few months before we saw some big strides school-wide.

Finally, the district's director of instructional technology commented on an informal survey of participating teachers:

> At the end of each year for the past two years now, I have given teachers a survey where they can tell what's on their mind about the project. The results have been amazing. This year, 97% of the teachers reported that their students are now more engaged and 83% of the teachers said that they are a better teacher now because of this project. This project has had a hugely positive impact.

Taken at face value, these quotes provide anecdotal evidence that the focused professional development had a positive influence on the teachers and students at the school. Although alternate explanations were not examined in the descriptive study, when the anecdotal evidence is combined with findings from the visual and statistical analyses, a reasonable inference can be made that the school's action research effort was a contributing factor to improved academic achievement at the school. This was accomplished through professional development that focused on using interactive technologies to enhance the instructional strategies articulated in the *Art and Science of Teaching* framework (Marzano, 2007). The professional development contributed to improving teacher proficiency with instructional strategies and interactive technology. Due to the descriptive nature of this case study, additional quantitative analysis is recommended to further examine the relationship between the focused professional development and student academic achievement.

Technical Notes

Technical Note 1

Pearson's chi-square test is commonly used to examine relationships between two categorical variables. For example, in education research, the chi-square test is often used to assess the relationship between the number of students who received an intervention and the number of students who passed a specific test. This type of data is typically displayed in a contingency table. To illustrate, Table TN1 shows the format for a 2 x 2 contingency table (two rows and two columns) that might be used for an education intervention program.

Table TN1: 2 x 2 Contingency Table Format

	Pass	Fail
Intervention		
No Intervention		

Pearson's chi-square (χ^2) is calculated with the following formula:

$$\chi^2 = \sum \frac{\left(observed_{ij} - model_{ij}\right)^2}{model_{ij}}$$

In the formula, *i* represents the rows in the contingency table of the data and *j* represents the columns. The observed data are the frequencies reported in each cell of the table. The model used in the formula is calculated from the expected frequencies with the following formula:

$$model_{ij} = E_{ij} = \frac{row\ total_i \times column\ total_j}{n}$$

In this example, expected frequencies would be calculated for each of the four cells of the table using the column and row totals for each cell to calculate the expected values:

$$\frac{RT_{intervention} \times CT_{pass}}{n}, \frac{RT_{intervention} \times CT_{fail}}{n}, \frac{RT_{no\ intervention} \times CT_{pass}}{n}, \frac{RT_{no\ intervention} \times CT_{fail}}{n}$$

The chi-square value is checked against a distribution with known properties. For a 2 x 2 table, any value greater than 3.84 would be considered statistically significant at the standard alpha of .05 (that is, $p < .05$). In other words, there would be less than a 5% chance of the chi-square test producing the same value or higher if there was actually no association between the variables. That is to say, there would likely be a relationship between the variables.

Technical Note 2

In social science research, null hypothesis significance testing (NHST) has been used for more than 50 years to make judgments about the results of statistical analyses (Harlow, 1997). When testing statistical hypotheses, researchers are often concerned with reaching incorrect conclusions. To address this issue, two rates of probability are typically considered, α, the Type I error rate (the probability of mistakenly rejecting the null hypothesis) and, β, the Type II error rate (the probability of mistakenly accepting the null hypothesis). By convention, the Type I error rate is typically kept at or below $\alpha = .05$ (Steiger & Fouladi, 1997). Researchers typically highlight results of analyses that meet this standard level of statistical significance (that is, $p < .05$), which is often interpreted as less than a 5% chance of mistakenly rejecting the null hypothesis. It is important to note that the Type I and Type II error rates represent conditional probabilities. Mulaik, Raju, and Harshman (1997) explained:

> . . . Type I and Type II error rates are never actual probabilities of making these errors. Having no knowledge about the true effects when setting out to perform a significance test, we have no way of knowing what the true error rates will be. So these error rates are hypothetical probabilities considered conditionally under the case where the null hypothesis is true and under a case where the null hypothesis is false, respectively. (p. 75)

In other words, at $\alpha = .05$ (that is, $p < .05$), the conditional probability of mistakenly rejecting the null hypothesis (assuming it to be true) would be less than 5%.

For example, an education researcher might be interested in conducting a simple two-group experiment to examine the influence of a recently developed math curriculum. In the experiment, one group of students would receive classroom instruction with the newly developed curriculum and the other group of students would receive classroom instruction without the new curriculum. Both groups of students would take the same math test and their scores would be compared with a statistical test, such as the independent-samples *t* test. Two statistical hypotheses might be tested, the null hypothesis of the mean (or average) scores of the two groups being equal ($H_0\colon \mu_1 = \mu_2$) and an alternative hypothesis of the mean scores of the two groups being different ($H_1\colon \mu_1 \neq \mu_2$).

The value obtained from the *t* test is compared with the maximum (or critical) value that would be expected by chance alone in a *t* distribution with the same degrees of freedom (Field, 2009). At $\alpha = .05$ (that is, $p < .05$), if the null hypothesis of no difference between the means of the two groups is assumed to be true, the conditional probability of getting a result as extreme or more extreme than the critical value would be less than 5%. In other words, the researcher can be 95% confident of the reported result. Conversely, a result that does not meet the standard level of significance (that is, $p \geq .05$) is often taken to suggest that the finding is more likely due to chance or random sources of error, with the likelihood increasing as the *p*-value approaches its upper limit of 1.00. For instance, when $p = .10$, the researcher can be 90% confident of the result, when $p = .20$ the researcher can be 80% confident of the result, and so on. (For a more detailed discussion of the arguments for and against the use of significance tests, see Harlow, Muliak, & Steiger, 1997.)

Technical Note 3

Osborne (2006) noted that the odds ratio effect size has a long tradition in biomedical research as a mechanism to examine contributing factors to disease or mortality. In recent years, logistic regression and odds ratios have become more common in social science research as a mechanism to predict categorical outcomes based on continuous or categorical predictor variables (Osborne, 2006). However, the odds ratio has been criticized as being less intuitive than the rate ratio (or relative risk) and therefore more easily misunderstood and misinterpreted (Osborne, 2006; see also Davies, Crombie, & Tavakoli, 1998; Grimes & Schulz, 2008; Sistrom & Garvan, 2004).

The odds ratio and rate ratio effect sizes are based on statistical probability and provide ways to examine the magnitude of difference between the likelihood of an event in two groups. Probability is often expressed as the following ratio:

$$P = \frac{\#\ of\ actual\ occurrences}{\#\ of\ possible\ occurrences}$$

For example, the probability of rolling a 2 on a six-sided die would be 17% (1 ÷ 6).

Odds are often expressed as a ratio of two probabilities (the probability of the event occurring to the probability of the event not occurring):

$$odds = \frac{P}{1 - P}$$

Considering the same example, the odds of rolling a 2 are approximately 20% (17% ÷ 83%). Therefore, the odds in favor of successfully rolling a 2 are 20:100 or 1:5 (one success to five failures). Stated in terms of failure, the odds against rolling a 2 are 5:1 (five failures to one success). In other words, rolling a number other than 2 would be five times more likely.

The odds ratio represents the magnitude of difference between the relative odds of an event in two groups and the rate ratio represents the magnitude of difference between the probabilities of an event occurring in each group. A value equal to 1 would indicate that there was no difference between the two groups, a value greater than 1 would indicate that there was a difference between odds (or probabilities) which favored the first group, and a value less than 1 would indicate that there was a difference between odds (or probabilities) which favored the second group.

Table TN2 displays a contingency table for a hypothetical educational program.

Table TN2: 2 x 2 Contingency Table for Hypothetical Educational Program

	Pass	Fail
Intervention	70	30
No Intervention	40	60

Considering the hypothetical data presented in Table TN2, the odds ratio effect size is calculated with the following formulas:

$$odds_{intervention} = \frac{\#\ of\ students\ who\ pass}{\#\ of\ students\ who\ did\ not\ pass} = \frac{70}{30}$$

$$odds_{no\ intervention} = \frac{\#\ of\ students\ who\ pass}{\#\ of\ students\ who\ did\ not\ pass} = \frac{40}{60}$$

$$odds\ ratio = \frac{odds_{intervention}}{odds_{no\ intervention}} = \frac{70/30}{40/60}$$

The odds of passing for students in the program are 70:30 or 7:3 in favor of passing. The odds of passing for students not in the program are 40:60 or 2:3 in favor of passing, or 3:2 against. The odds ratio for this hypothetical program is 3.48 (2.33 ÷ 0.67). The odds ratio effect size indicates that the odds of passing the test were 3.48 times higher if students received the intervention than if they did not.

The rate ratio effect size is calculated with the following formulas:

$$P_{intervention} = \frac{\#\ of\ students\ who\ pass}{\#\ of\ students\ who\ receive\ intervention} = \frac{70}{100}$$

$$P_{no\ intervention} = \frac{\#\ of\ students\ who\ pass}{\#\ of\ students\ who\ did\ not\ receive\ intervention} = \frac{40}{100}$$

$$rate\ ratio = \frac{P_{intervention}}{P_{no\ intervention}} = \frac{70/100}{40/100}$$

The probability of passing for students in the program is 70% and the probability of passing for students not in the program is 40%. The rate ratio for the fictitious program is 1.75 (70% ÷ 40%). The rate ratio effect size indicates that the probability of passing the test is 1.75 times higher if students received the intervention than if they did not. In other words, students in the program would be 75% more likely to pass the test than students not in the program.

References

California Department of Education. (2013a). *DataQuest*. Retrieved from http://dq.cde.ca.gov /dataquest/

California Department of Education. (2013b). *2012–13 academic performance index reports: Information guide*. Retrieved from http://www.cde.ca.gov/ta/ac/ap/

Comprehensive Meta-Analysis (Version 2.2.057) [Computer software]. Englewood, NJ: Biostat.

Davies, H. T. O., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *British Medical Journal*, *316*, 989–991. doi:10.1136/bmj.316.7136.989

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Grimes, D. A., & Schulz, K. F. (2008). Making sense of odds and odds ratios. *Obstetrics & Gynecology, 111*(2), 423–426. doi:10.1097/01.AOG.0000297304.32187.5d

Harlow, L. L. (1997). Significance testing introduction and overview. In L. Harlow, S. Mulaik, & J. Steiger (Eds.), *What if there were no significance tests?* (pp. 1–17). Mahwah, NJ: Erlbaum.

Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (1997). *What if there were no significance tests?* Mahwah, NJ: Erlbaum.

IBM SPSS Statistics (Release 22.0.0.0) [Computer software]. Armonk, NY: IBM Corp.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Marzano, R. J. (2007). *The art and science of teaching: A comprehensive framework for effective instruction.* Alexandria, VA: ASCD.

Marzano, R. J., & Haystead, M. W. (2009). *Final report: Evaluation study of the effects of Promethean ActivClassroom on student achievement*. Centennial, CO: Marzano Research Laboratory.

Marzano, R. J., & Haystead, M. W. (2010). *Final report: A second year evaluation study of Promethean ActivClassroom*. Centennial, CO: Marzano Research Laboratory.

Mulaik, S. A., Raju, N. S., & Harshman, R. A. (1997). There is a time and place for significance testing. In L. Harlow, S. Mulaik, & J. Steiger (Eds.), *What if there were no significance tests?* (pp. 65–115). Mahwah, NJ: Erlbaum.

Osborne, J. W. (2006). Bringing balance and technical accuracy to reporting odds ratios and the results of logistic regression analysis. *Practical Assessment Research & Evaluation*, *11*(7). Retrieved from http://pareonline.net

Richtel, M. (2011, September 3). In classroom of future, stagnant scores. *The New York Times.* Retrieved from http://www.nytimes.com

Sistrom, C. L., & Garvan, C. W. (2004). Proportions, odds, and risk. *Radiology*, *230*, 12–19. doi:10.1148/radiol.2301031028

Steiger, J. H., & Fouladi, R. T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. Harlow, S. Mulaik, & J. Steiger (Eds.), *What if there were no significance tests?* (pp. 221–257). Mahwah, NJ: Erlbaum.