PPIC PUBLIC POLICY INSTITUTE OF CALIFORNIA

Technical Appendix

Lessons in Reading Reform

Finding What Works

Julian R. Betts, Andrew C. Zau, and Cory Koedel

Contents

Introduction
Background and Data
Effects of the Interventions on Reading Test Scores
Effects of the Inverventions on Math Test Scores, Absences, and Grade Retention
Outcomes at the High School Level
Conclusion
Tables 1–19
References

Supported with funding from the Donald Bren Foundation

Introduction

This technical appendix provides more detail on the reading reforms implemented under the Blueprint for Student Success project in the San Diego Unified School District (SDUSD) between 2000 and 2005. It provides details on the dataset, the econometric methods we employ, and the results, which are also detailed and discussed in the main report.

From a policy perspective, it is imperative to learn whether the reforms in San Diego boosted achievement. From a statistical standpoint, the reforms also offer a relatively clean way to identify the effect of the reforms, compared to the traditional literature on education production functions. The latter literature, which began in earnest with Coleman (1966), typically uses variations in class size and teacher qualifications across students that are potentially endogenous.

More recent variants in this literature have improved identification by using student panel data combined with student fixed effects, thus using the within-student variation to identify the effects of variation in school and classroom resources on student outcomes. We adopted this approach. Nonetheless, a concern in the fixed-effects work is that variations from year to year in a student's classroom environment reflect unobserved variations in student performance or motivation.

The San Diego setting is particularly amenable to the use of student fixed effects. With one minor exception, this is due to the elements of the Blueprint being implemented in the summer of 2000. Our data panel, which begins in 1999, facilitates a before-and-after analysis. In addition, the district curtailed and in some cases eliminated some of the reform elements starting in the 2003-2004 school year. This time pattern of implementation and subsequent winding down of the reforms, plus variations in which reform elements were available for a student as he or she progressed from grade to grade, creates a largely exogenous source of variation in exposure to the reforms, thus enhancing the credibility of the student-fixed-effect formulation.

Background and Data

Several studies have previously analyzed the Blueprint for Student Success. The American Institutes for Research (AIR) was hired by the SDUSD school board to produce evaluations of the first two years of the Blueprint (2000-2001 and 2001-2002). See AIR (2002) and Quick et al. (2003).[1] The AIR research compared progress in student achievement, measured using several of the statewide tests, between SDUSD and a number of comparison districts. Bassok and Raymond (2005) produced a similar follow-up analysis. While useful, these aggregate studies cannot tell us whether the reforms worked in an absolute sense, because the comparison districts in California were implementing reading reforms at the same time. In particular, the Los Angeles Unified School District patterned many of its reforms on the San Diego precedent. More fundamentally, these analyses of aggregate trends in test scores cannot attribute changes in average achievement to the Blueprint specifically, nor can they identify which of the many interventions were the most or least effective.

---

[1] See also Stein, Hubbard, and Mehan (2004) for a perceptive comparison of the "cultures of reform" in SDUSD and New York City's District #2. Former SDUSD Chancellor of Instruction Tony Alvarado made major contributions to academic reforms in both districts. Hightower (2002) and Betts (2005) also provide an overview of the reforms.

The present paper builds directly on the work of Betts, Zau, and King (2005), the first student-level analysis of the impact of the reforms. These authors found that the reforms had a moderate positive effect on achievement in elementary schools, a small positive effect in middle schools, and a moderate negative effect in high schools. They also found that some of the student interventions were quite effective and that peer coach system had yet to bear significant fruit. A limitation of that study is that it analyzes the reforms only through spring 2002, the end of the second year of the reform. It is important to know whether early gains were sustained, diminished over time or grew as district teachers and administrators gained experience with the reforms. The present study also extends the previous work by examining a far richer array of outcomes, including the probabilities of grade retention, the completion of the college-preparatory course sequence in high school, and completion of high school.

Data

Our data-set consists of complete student academic records, including test scores, courses taken, and absences, from fall 1999 through spring 2005. The data include indicators of Blueprint interventions in which each student participated in a given year, and a rich set of covariates related to the student's class size and teacher qualifications (overall in elementary school, and in English in middle and high school).

California has administered different tests at different times. It mandated the Stanford 9 test in spring 1998 through spring 2002, and the California Standards Test (CST) in spring 2002 and later years. We convert both into z-scores by subtracting the district-wide mean and dividing by the district-wide standard deviation for a given year and grade. In spring 2002 we use the CST. Because the state gave both the Stanford 9 and CST tests in that year, we can examine the degree to which individual students' results are consistent across tests. In reading, the main focus of this paper, the correlation across tests was 0.93, and in math it was 0.86. These are extraordinarily high; they considerably exceed the correlations of CST scores in 2002 and 2003 for individual students, 0.85 for reading and 0.76 for math. The high correlations are in part by design because the designers of the CST borrowed many questions from the Stanford 9.

Blueprint program elements, 2000-2001

	Blueprint element	**Student group**	**Content**
Prevention	Genre Studies	Students reading at or above grade level, grade 6 or 7	Two-period English class
	Peer coaching	All students	Placed at all schools for teacher development
	Focus schools	All students in the lowest decile of elementary schools	Extended year, additional peer coach, additional funding
	API-2 schools	All students in the second-lowest decile of elementary schools	Additional peer coach, additional funding
Intervention	Literacy Block	Students reading below grade level, grades 6–10	Double-length English classes
	Literacy Core	Students significantly below grade level in grade 9	Triple-length English classes
	Extended Day Reading Program (EDRP)	Students below and significantly below grade level in all schools (grades 1 through 9)	Three 90-minute periods of supervised reading each week before or after school
	Summer school/ intersession	Students in most grades from kindergarten through grade 9 who lagged in reading	Six weeks, four hours per day, of reading in summer, or intersession

Effects of the Interventions on Reading Test Scores

We model gains in test scores, or $\Delta Score_{icgst}$ for student i in classroom c in grade g in school s in year t as a function of school, family and personal, and classroom characteristics. Our regression model is:

$$\Delta Score_{icgst} = \alpha_s + \beta_{Zipcode_{it}} + \gamma_i + Score_{icgs,t-1}\omega + \mathbf{FAMILY}_{it}\mathrm{E} + \mathbf{PERSONAL}_{it}\Phi + \mathbf{CLASS}_{icgst}\Gamma + \mathbf{SCHOOL}_{ist}\Lambda + \mathbf{BLUEP}_{it}\mathrm{K} + PEER_{igs,t}\pi + \varepsilon_{it} \quad (1)$$

where the first three regressors represent fixed effects for the student's school, home zip code, and also the student; $Score_{icgs,t-1}$ is the student's prior year score, added as a control for regression to the mean; the next four items in bold characters indicate vectors of time-varying family, personal, classroom, and school characteristics (including year dummies); $\mathbf{BLUEP}_{ist}$ is a vector characterizing student i's participation in Blueprint interventions in year t, along with measures at the school level of Blueprint elements such as peer coach to enrollment ratios expressed as a percentage; $PEER_{igs,t}$ is the average test scores of a students' peers in his or her grade level at the current school, based on the prior spring's tests and expressed as the number of standard deviations above or below the district mean score for that grade and year; corresponding Greek letters are vectors of coefficients and ε_{it} is an error term.

We include a lagged score because of evidence that there is regression to the mean in the test-score data. Regression to the mean is problematic because a student who has had a bad year academically is likely to have a larger-than-usual gain in test scores the next year and also to participate in a Blueprint intervention. If we do not control for regression to the mean we could wrongly attribute the abnormally large gain in test scores to the intervention itself. An analogous problem was first noted in the literature on government training programs by Ashenfelter (1978). Below, we discuss how results change when we re-estimate (1) without the lagged test score. In some cases we find larger (more positive) coefficients on Blueprint variables, which is consistent with this problem (now commonly known as "Ashenfelter's dip").

Although we include student fixed effects to account for any unchanging and unobserved aspects of students, schools, and zip codes, there are many confounding factors related to students, their families, and their schools and classrooms that could change over time. We incorporate these as explanatory variables. We do not highlight the effect of these explanatory variables on reading achievement in this paper, but it is nonetheless important to control for them.[2] Table 1 lists all of the explanatory variables that we add to the model of gains in reading scores for elementary school students. These variables include aspects of students, families, and neighborhoods that could change over time. Table 2 lists explanatory variables at the school, student body, grade, and classroom level that we use in our main models for elementary schools. They are mostly self-explanatory, with the possible exception of some of the teacher characteristics. We include controls for teachers with a full credential and two types of teachers with less than a full credential—those with an emergency credential, and interns. We also interact these variables with dummy variables for the teacher's total years of teaching experience (0–2, 3–5, and 6–9). We also control for a number of language certifications that certify that a teacher has received training in how to teach English Learner (EL) students. The first, CLAD, (the Crosscultural, Language and Academic Development credential), prepares teachers to

[2] See Betts, Zau, and Rice (2003) for a detailed accounting of the influence of these other variables on gains in reading achievement. The results reported there are quite similar to what we found in the present models.

teach EL students. BCLAD, the Bilingual CLAD, is similar but prepares bilingual teachers to teach in a bilingual classroom.

At the middle school and high school level we include all of the explanatory variables listed in Tables 1 and 2 with three modifications. First, whereas we focus on each elementary student's homeroom teacher, at the middle and high school levels we instead focus on the characteristics of each student's English classroom and English teacher. Second, at the middle school and high school levels we need to control for additional characteristics of teachers, including subject authorizations, which indicate the degree of mastery of the subject matter at hand. A teaching credential, on the other hand, denotes mastery of more general approaches to teaching. Subject authorization levels include—in descending order—full authorization, supplementary, board resolution, and limited assignment emergency (LAE).[3] Accordingly, we add controls for a supplementary, board resolution, or LAE subject authorization. Third, at the middle school and especially the high school level, the number of English classes that a student takes each year may vary. We therefore add indicator variables showing that students took zero or one English course, or more than (the normal load of) two classes in a given year.

To these models we added numerous characterizations of Blueprint elements. Peer coaches are placed in schools to assist classroom teachers by observing their lectures and providing feedback, providing lectures while the regular teacher watches, and providing training in various other ways. We wanted to test whether the intensity of peer coaches in a school influenced reading gains. Therefore we calculated the ratio of coaches to overall enrollment in each school, expressed as a percentage (that is, multiplied by 100). Because class size varies little across schools in the district (Betts, Zau, and Rice, 2003), an individual coach who had to work with a greater number of classrooms could be less effective.[4] Because a coach's own experience might influence his or her effectiveness, we also included a measure of the average years of teaching experience of peer coaches at the school.

At the elementary school level, two important Blueprint elements are the Focus and API-2 schools, which receive substantial additional resources. We add indicator variables for which elementary schools were in these groups. We note that one of the additional resources that schools in both categories received was a second peer coach. Because we control separately for this, we can effectively distinguish between the effect of peer coaches on the one hand, and on the other, the collective effect of the other resources added through the Focus and API-2 school programs.

Because the student fixed-effects formulation removes variation across students, in effect mean-differencing the data, identification of the Blueprint variables comes from within-student variation in exposure to various aspects of the reform. A major source of within-student variation comes from the way in which the program was implemented and then, over the course of several years, pared back (mainly due to financial realities). In 1999-2000, the first year for which we have a test-score gain, the Blueprint was not in place, except for peer coaching and Genre Studies, which were implemented on an extremely limited basis, and Literacy Block,

---

[3] Full and supplementary subject authorizations are official authorizations mandated by the California Commission on Teacher Credentialing (CCTC). Board resolutions refer to decisions by the SDUSD board to authorize a teacher to teach a specific subject, based on the teacher having taken relevant college courses. These teachers may lack one or two courses required for a supplementary authorization or have enough in the general subject area but not the exact set of courses required by the CCTC. LAE authorizations are short-term authorizations for teachers with less subject knowledge. These should not be confused with emergency credentials, because LAE credentials are given to fully credentialed teachers teaching outside of their normal assignment. Some high school teachers may not hold any of the above subject authorizations, because they are not yet fully credentialed teachers.

[4] In early versions of this work we also tried a regressor that simply counted the number of peer coaches at the school. The results were qualitatively similar.

which was implemented in grades 9 and 10 in 1999-2000. Most elements of the reform were implemented in 2000-2001 and expanded in 2001-2002. (The preventive Genre Studies class did not fit this pattern exactly because it was first introduced in grade 6 and then, in 2002-2003, in grades 7 and 8.) Many programs were scaled back or eliminated around 2003-2004 or 2004-2005. This provides a second source of within-student identification. Table 3 shows the percentage of students in our sample who participated in the various literacy interventions by year. To save space we do not show a column for 1999-2000, except for programs which that year had participation above zero.

Main Results

Table 4 shows the main results for reading at the elementary school level. It would be misleading to attribute to the literacy reforms any variation in achievement gains that in fact arose due to variations in class size or teacher qualifications. Model i excludes controls for class size and teacher characteristics. Models ii and iii add controls for class size and teacher characteristics respectively, while model iv includes controls for both class size and teacher characteristics. The coefficients and level of significance of the Blueprint variables are affected little by the addition of these controls. (The main report summarizes results from model iv.) Here, we report on models of gains in z-scores, so that that the interpretation of the regression coefficients is that a one-unit increase in the given regressor is predicted to change the student's test score by the corresponding number of standard deviations of test scores indicated by the coefficient. While it is common in the academic literature to present results this way, for our report we converted these estimates to the number of percentile points a student would move up or down in the rankings of students, in each grade, after participating in a given Blueprint element. To make this conversion, we assumed that most Blueprint elements were applied to a student at the 25th percentile of the district distribution, because the interventions were targeted at students lagging behind in reading. Thus, we added the coefficient to the value of a standard normal at the 25th percentile and calculated the change in the percentile directly from the cumulative distribution for a standard normal distribution. In addition, the Focus school and API-2 preventive measures were targeted at the lowest-performing schools and were therefore treated in the same way. There were two exceptions to using the 25th percentile as the base. Peer coaching was a form of professional development provided for all teachers (and implicitly, designed to help all students). Thus we converted the estimated effects assuming that the student began at the 50th percentile. The second exception was Genre Studies, the preventive-strategy double-length English classes aimed at students reading at or above grade level. We converted the estimated effects assuming that students in Genre Studies were initially at the 75th percentile.

The measure of peer coaches as a percentage of enrollment has a positive coefficient but is not quite significant. A second variable indicating the average years of teaching experience of peer coaches is negative and insignificant. Two other preventive measures, the Focus and API-2 schools, both enter in a positive and significant fashion. The coefficients are small, at roughly 0.02 and 0.03 respectively. (Because the dependent variable is gains scaled in terms of district-wide standard deviations within grade, we can interpret these coefficients as the predicted gain in standard deviations resulting from a student's school receiving the district supports under these programs.) EDRP and Blueprint summer school elements are not significant, but the intersession literary program (designed for students at year-round schools for whom summer school is impractical) enters in a positive and significant fashion.

A number of value-added test-score studies have found evidence that teachers in their first few years of classroom experience do not teach as effectively as do more experienced teachers. It stands to reason that the effect of Blueprint interventions might vary with the experience of the homeroom teacher. The direction of the effect is unclear a priori. Blueprint interventions might be more effective when the teacher is inexperienced if they act as a substitute for teacher experience. Conversely, if they act as a complement to teacher experience, then the interventions could be more effective when the homeroom teacher is more experienced.

Model v adds interactions between dummies for teacher experience and the Blueprint elements that take place during the regular school day: peer coaching, and the Focus and API-2 school programs. An F-test (bottom row of the table) suggests that the interactions belong in the model. Two interactions, with the peer-coach-to-enrollment variable and the indicator for API-2 schools, enter negatively and significantly, suggesting that students with more highly experienced teachers gain more from the peer coaching and API-2 programs. The main peer-coach-to-enrollment variable also becomes significant and positive in this model. However, this main effect, with a coefficient of 0.04, does not suggest a large effect of peer coaches. As shown in Table 14, a one standard-deviation increase in the peer coach to enrollment percentage is 0.13; an increase of that amount is predicted to increase achievement by 0.005 of a standard deviation.

Model vi further adds interactions between the homeroom teacher's experience and EDRP and Summer session Blueprint elements. In a few instances we find negative interactions between the intervention and dummies indicating relatively less experienced homeroom teachers. The interaction effects, like the main effects, are small.

Table 5 shows results for middle schools. We focus on model iv because the results do not vary much with the addition of variables for class size and teacher characteristics. Indicators for the double- and triple-length classes for students below grade level in reading, the Literacy Block and Literacy Core Blueprint elements, are both positive and highly significant. The latter coefficient is meaningful, suggesting one year in Literacy Core is associated with a boost of 0.16 standard deviations in achievement. The Genre Studies variable also enters positively and significantly. These double-length English classes for students who were at or above grade level appear to have about the same influence on gains in test scores as does Literacy Block, the corresponding classes for students the district believed to be reading below grade level. In addition, we include a control for the Accelerated class element, which was grade retention followed by triple-length English classes in the year in which the student was retained. The coefficient on this variable is negative and significant, but small. The remaining Blueprint variables do not enter significantly.

Models v and vi add interactions between selected Blueprint variables and indicators for the English teacher's experience. As shown at the bottom of Table 5, the null of no interactions is handily retained. A few of the interactions are significant, providing some evidence that Genre Studies and EDRP could be relatively less effective when the English teacher is relatively inexperienced. These effects are very small.

At the high school level, as shown in Table 6, all of the Blueprint variables are either insignificant or enter significantly but negatively. The latter applies to Literacy Block, Literacy Core, Summer session, and average years of teaching experience of peer coaches. Betts, Zau, and King (2005) report similar results using a much shorter panel. No evidence of significant interactions with the level of experience of the student's English teacher emerges in models v and vi.

Other Specifications

As mentioned earlier, because of regression to the mean in test scores, we felt it important to control for the initial test score so as to avoid attributing students' natural improvement after a randomly bad year to their subsequent placement in the Blueprint reforms. Table 15 reproduces model iv from Tables 4 through 6 for each grade span and then repeats these models but without the lagged reading score. In general we find that omitting initial achievement leads to larger (more positive) coefficients on the Blueprint variables. In some cases the changes are meaningful. In the models for elementary schools, the estimated effect of EDRP participation becomes statistically significant, positive, and fairly large. In the middle school model, five of the eight Blueprint variables either become positive and significant, or, if they were already positive and significant, the coefficients become larger. Of the remaining three variables, two remain statistically insignificant and one, Genre Studies, which is positive and significant in the model with the lagged score, becomes insignificant in the model without the lagged score. Perhaps the most notable findings are that EDRP and Summer session become positive and significant, although with modest effect sizes of about 0.05. At the high school level, Literacy Block and, for EL students, Literacy Block/Core, switch from being negative and significant to positive and significant, but the coefficients on the other variables change in less dramatic ways.

On the whole, then, the results without the lagged scores are far more favorable to the Blueprint than our main specification, but it seems likely that much of this is due to students having a bad year, being placed in a Blueprint intervention the next year, and then naturally regressing to the mean (that is, having a larger-than-usual gain in test scores). Given our concern about regression to the mean, and the evidence in this table that regression to the mean may materially alter the conclusions, we restrict our attention to our original model (1), in the knowledge that the findings from this specification are conservative.

Next, we tested for variations in the effect of various Blueprint elements by year. We anticipated several possible patterns. One is increased effects of each Blueprint element each year. Such patterns would emerge if there were learning-by-doing effects as the interventions were rolled out. A second possibility is a pyramid or inverted-U pattern of effectiveness against time. One explanation would be initial increases in effectiveness followed by a period of reversals. There are two reasons for the effectiveness of the reforms to decline eventually. First, union and teacher resistance to the reforms was quite strong, and appeared to grow over time. In addition, state cutbacks to public education led to budget shortfalls in districts around the state. A third possible pattern that could emerge is one that would be consistent with Hawthorne effects, namely an immediate positive effect that diminishes very quickly as the psychological effect of a school (or a set of underperforming students) receiving newfound attention wears off.

Tables 16, 17, and 18 show the results of re-estimating model iv after adding year interactions with various Blueprint interventions. The tables show the main effect of each policy for 2000-2001 as well as interaction terms between Blueprint variables with dummy variables for other years. The second column in these tables indicates whether the total effect of the given policy in a given year (which will be the sum of the main effect and, for all but the base year, the interaction term for the given year), is statistically significant.

No single pattern emerges as dominant in these models. The most typical pattern is the pyramid, in which effectiveness of a given Blueprint element increased for one or more years and then began to fade. Examples of this include the Focus and API-2 programs in elementary schools, Genre Studies in middle schools, and

Literacy Core for non-EL students in high school.[5] The Focus effects exhibit a particularly steep rise and decline. Focus schools had a longer school year (about 24 days) only in 2000–2001 and 2001–2002, after which financial constraints forced the cancellation of the longer school year. This pattern coincides with higher reading test scores in spring 2002 and spring 2003. A further indication that the longer school year may have been crucial is the fact that only in these two years is the estimated effect of the Focus program statistically different from zero.[6]

A second common pattern was gradual increases in effectiveness over time. It is striking to see that in middle schools, the coefficient on Literacy Block/Core for EL students becomes bigger each year, without exception. In high schools, the effects of Literacy Block/Core for EL students also increased (in the sense that they became less negative each year, before becoming positive but not significant in 2004-2005). The implication is that middle and high school teachers (and their EL students) made better use of these extended-length English classes each year.

Our measure of the peer coach element as a percentage of enrollment shows similar increases in effectiveness over time, although the pattern is not completely uniform. This makes perfect sense because it should take some time for coaches to begin to visit with classroom teachers, to model teaching methods, and to encourage adoption and mastery of the methods by the classroom teachers.

The effects of Blueprint elements may have varied over time in other ways. In our test score models, we assume that an intervention in grade 8 affects achievement gains in grade 8 but not in later grades. It is possible that such gains are temporary, so that larger-than-average gains for students in grade 8 would be followed by smaller-than-average gains in grade 9. Conversely, it could be that participating in a reading intervention in grade 8 boosts achievement gains in grade 8 as well as in grade 9 and beyond. Data availability limits our ability to test these possibilities: We have a fairly short panel, and in addition, there is some collinearity between participating in a given Blueprint element in year t and year t-1. When we repeated our test-score models adding the previous year's Blueprint participation, by far the most common finding was that exposure to a Blueprint element in the prior school year increased achievement gains in the current year. This occurred in just over half the cases. In only about 10 percent of cases there was evidence of a negative effect in a later school year, and these cases involved high school interventions that we had already estimated to have an overall negative effect. In the remaining 40 percent of cases, no evidence of an effect of past exposure to a Blueprint element emerged. Our interpretation of these findings is somewhat limited because of the collinearity in the data, which in a few cases manifested itself by the current-year effect becoming insignificant once we added controls for the same Blueprint element in the prior year. We tentatively conclude that, in elementary and middle schools, in many cases, the Blueprint led to both immediate as well as perhaps future gains in achievement; in high schools, the negative effects sometimes spilled over into the following year.

[5] The last of these is slightly different in that the overall effect never becomes positive and significant. But the same sort of pyramid or inverted-U pattern persists: the coefficient is negative and insignificant in all years except for 2002-2003, when it is negative but closest to zero, and not significant.

[6] We thank Karen Bachofer for this insight.

Effects of the Interventions on Math Test Scores, Absences, and Grade Retention

In this section we assess, on an annual basis, possible effects of the Blueprint on outcomes apart from reading achievement. Early opponents of the Blueprint, including a Latino coalition in San Diego, expressed concern that the Blueprint, by placing such emphasis on literacy, would distract students from other subjects. We call this the academic diversion hypothesis. Countering this idea is one that posits reading as a gateway skill, mastery of which allows a student to learn better from textbooks in all subject areas. A second hypothesis of opponents was that students would burn out, especially from the extended literacy classes and perhaps EDRP, the after- and before-school reading program. One way to assess this theory is to model student absences. Another way is to ask whether participants were more likely to be retained a grade in years in which they participated in various interventions. (Recall that the Blueprint called for grade retention in certain grades, but this was never implemented in a meaningful way apart from middle school, and even there, only on an extremely limited basis.)

Table 7 models gains in math achievement for students, using the same specification as in model (iv) of Tables 4 to 6, except that now the lagged score is in math, and in the middle and high school models, classroom and teacher characteristics refer to the math classroom and teacher (not the English classroom and teacher). A few of the Blueprint variables are significant in each of the models for the various grade spans, but the coefficients are usually very small, and they are roughly balanced between positive and negative effects. (The largest negative coefficients obtain in high school, suggesting that participation in Literacy Block or Literacy Core is associated with a drop in the annual gains in math achievement of about 1/20 of a standard deviation.) We conclude that at least by this measure, Blueprint interventions did not divert students' attention strongly away from learning math.

Table 8 models the percentage of the school year a student is absent. Here the evidence is again mixed, but overall more supportive of the view that the literacy programs did not systematically cause burnout. In both elementary and high schools, there were three cases of Blueprint elements that were negatively associated with student absences, and only one element that was positively associated. Moreover, the negative coefficients tend to be larger than the positive coefficients.[7] But in middle schools, five Blueprint elements were found to be positively associated with absences, compared to just three that had a negative association. Most notable here are the coefficients related to Literacy Block and Literacy Core, participation in which is predicted to increase absences by about 0.3 and 0.4 percentage points for non-EL and EL students respectively. At the high school level participation in these programs is predicted to actually decrease absences by about 0.4 percentage points.

Table 9 presents versions of model iv that replace the gain in test score with a dummy indicating whether the student was grade-retained, that is, held back in a given year. In the elementary and middle school models, a few Blueprint variables enter significantly, in varying directions, and the coefficients are quite small. In high schools, the predicted effects are again small, but they are uniformly negative, suggesting that participation did lower the probability of being retained. The largest coefficient in the table, in absolute value, is -0.036 for

---

[7] This point is hard to see for the variable "peer coaches as a percentage of enrollment," which has a positive coefficient in elementary schools and a much larger negative coefficient in the high school model. Based on the descriptive data in Appendix Table 1, a one-standard-deviation increase in this variable is predicted to increase absences in elementary school by 0.06 percentage points and to lower absences in high school by an identical amount. In middle schools, where the effect is only marginally significant, the predicted changes in absences is +0.04 percentage points.

Literacy Core in high school, thus implying a drop of just under 4 percent in the probability of being retained for those participating in this element. In models not shown, we also replaced all of the variables capturing participation in the various Blueprint elements with a single variable indicating the number of interventions in which the student participated in the given year. At the elementary level, the coefficient was significant at 1 percent, equaling -0.001; at the middle school level the variable was insignificant; at the high school level the coefficient was again significant at 1 percent, equaling -0.009. Our overall conclusion is that the link between Blueprint participation and grade retention is weak and typically small and negative.

Outcomes at the High School Level: Graduation and College Admission Course Requirement Completion

In this section we evaluate the extent to which Blueprint interventions affect whether students graduate from high school and complete the requirements (the A–G high school course sequence) for admission to California's two state higher-education systems. These both represent longer-term outcomes than we have modeled above. Because we observe these outcomes only once, we cannot use our previous technique of student fixed effects to remove ability bias. As argued in the earlier section on test scores, we expect that students are negatively selected into the Blueprint program. This will limit our ability to make causal inference: The coefficient on Blueprint participation should be negatively biased in models of either graduation or completion of the A–G requirements because, by definition, only the weakest students participate in the Blueprint program.

As evidence that those who participate in Blueprint interventions are relatively academically challenged, we repeated our test score models without a student fixed effect. In this less-than-ideal approach, we compared one student to another rather than comparing the same student's achievement gains in different years, with and without Blueprint participation. In most cases, the estimated effect of Blueprint elements became smaller. For instance, in middle school, the estimated effect of Literacy Block switched from being positive to being negative, and increased in magnitude by about 50 percent. Instead of having zero effect on reading gains, the elementary Summer school intervention is estimated to have a large negative effect. These changes are almost surely due to the inability of these models to fully account for differences among students. Similarly, the results for graduation and A–G course completion are likely to be overly pessimistic because we are compelled to compare one student to another.

We start by evaluating students' graduation outcomes. Because graduation outcomes are only observed one time for each student, in lieu of a student fixed effect, we use the rich set of variables available in our dataset to remove as much of the negative selection bias as possible. Our controls closely match those listed in Tables 1 and 2 for the earlier models of test scores. At the student level, these variables include indicators for race, gender, English-learner status, parental education levels, and students' standardized test-scores in math and reading (Stanford 9) at the end of grade 8. At the school level we include school and zip code fixed effects; grade level peer performance; controls for the makeup of each school in terms of the percentage of students who are eligible for reduced-price lunch, are English learners, are new to the district, are new to the school or who changed schools that year; each school's racial composition; and each student's average class sizes in math and English during high school. We also control for the average qualifications of each student's teachers (in math and English) for experience, credentials, education levels, and the subjects of their bachelor's degrees. Where relevant, these variables are averaged within students over the course of their entire high school careers.

The first three columns of Table 10 show differences in means, both conditional and unconditional, and between Blueprint participants and non-participants in terms of graduation outcomes (estimated via linear probability models). Not surprisingly, students who are involved in the Blueprint project are much less likely to graduate from high school even when controlling for our substantially rich set of student- and school-level covariates. The first column, which does not include covariates, suggests that students who have participated in at least one Blueprint intervention are about 9 percent less likely to graduate from high school. This is clearly not causal. This coefficient reverses sign and becomes weakly significant once we control for student characteristics, but in the third column, where we add school characteristics, the coefficient again becomes negative and significant, but small. If taken literally, the third and fullest specification in this panel suggests that participating in the Blueprint is associated with a drop of 0.02, or 2 percent, in the probability of graduating from high school. This is a notable drop, given that approximately 89 percent of the student sample ultimately graduated from high school and just over 29 percent participated in at least one Blueprint intervention. Still, this effect is only one quarter as large as that of the model in the first column that fails to control for student characteristics.

If our first set of estimates in Table 10 were free from selection bias, they would imply that Blueprint interventions actually cause students to drop out of school. However, despite our efforts to control for selection bias, these estimates are unlikely to be completely unbiased by negative student selection.

We can somewhat mitigate the selection bias in our estimates by evaluating the effects of additional Blueprint interventions conditional on participating in at least one intervention. If the majority of the negative selection bias is associated with participation, and not additional interventions among participants, we can provide estimates of the effects of additional interventions that are relatively unbiased. In the final three columns of Table 10 we display the estimated effects of additional Blueprint interventions conditional on having ever participated in Blueprint. Despite the potential for some negative selection bias to remain in our estimates (students who receive the most interventions are likely to be the most needy), the results in the final column indicate that additional Blueprint interventions reduce dropout rates among participants on the order of about 2 percent per additional intervention. Countering this is a negative coefficient of about -0.06 on the Blueprint participation indicator. Together, the results in the final column suggest a negative correlation between Blueprint participation for those who have taken part in one or two interventions, roughly a zero effect for those who participated in three interventions, and a positive effect for those who took part in four or more interventions.

Like our graduation analysis, we evaluate the effects of Blueprint interventions on whether students complete the A–G course requirements. This issue is of particular concern because it has been argued that the additional hours in the classroom that are imposed by a Blueprint element take time away from underperforming students–time that they need to complete the A–G sequence. Again, we rely on the same set of student- and school-level variables to remove as much bias as possible from negative selection into the Blueprint program.

The first three columns of Table 11 show differences in means, both conditional and unconditional, between Blueprint participants and non-participants in terms of the completion of the required courses (estimated via linear probability models). Approximately 36 percent of our student sample completed the course requirements and again, 29 percent participated in at least one Blueprint intervention. The first model does not condition on anything apart from Blueprint participation. It reveals that Blueprint participants were, overall, 27 percent less likely to complete the A–G course sequence than other students. Again, this is clearly not causal. Just as with the preceding analysis of high school graduation, when we add controls for student and school characteristics, the coefficient becomes far smaller. If taken at face value, the third model, which most fully controls for initial student and school characteristics, implies that participation in the Blueprint is associated with a reduction of about 8 percent in the probability of completing the A–G course sequence.

Our initial analysis shows a strong negative correlation between Blueprint participation and the completion of the A–G sequence. In the final three columns of Table 11, we evaluate the effects of yet more Blueprint interventions on the completion of the course requirements, conditional on having ever participated. Our estimates indicate that additional Blueprint interventions apparently do not affect whether students complete the course requirements, at least when we control both for student and school characteristics. This result is somewhat surprising because additional Blueprint interventions are unambiguously time-consuming. Because students are time-constrained in terms of the number of classes they can take during high school, we might expect Blueprint interventions to limit students' capacities to complete other classes. On the other hand, to the extent that Blueprint participation improves a student's literacy it could increase his or her capacity to enroll in college preparatory courses later in high school. Importantly, to the extent that our control for having ever participated in the Blueprint project is not fully absorbing the negative selection effect, estimates in the final three columns in Table 11 will be biased downward.

In addition to our baseline results above, the A–G course completion analysis offers an opportunity to evaluate the effects of Blueprint interventions relatively free from any selection bias from either observed or unobserved student characteristics. Specifically, we can exploit our panel dataset to examine how Blueprint interventions affect class-taking behavior on a year-by-year basis for each student using the familiar within-student approach that was implemented in our test-score analysis. We estimate regressions of the following form:

$$\begin{aligned} Y_{igst} &= \alpha_s + \beta_{Zipcode_{it}} + \gamma_i + \mathbf{FAMILY}_{it}\mathrm{E} + \mathbf{PERSONAL}_{it}\Phi \\ &+ \mathbf{SCHOOL}_{ist}\Lambda + \mathbf{BLUEP}_{it}\mathrm{K} + PEER_{igs,t}\pi + \varepsilon_{it} \end{aligned} \quad (2)$$

Y_{igst} indicates how many university-certified classes are taken by student *i* in grade *g* in school *s* in year *t*, and other variables are as listed in (1), except that $BLUEP_{it}$ is now either a vector of participation indicators for Blueprint interventions, or instead, simply consists of the number of Blueprint interventions participated in by student *i* in the given year.[8] Our ability to control for unobserved student heterogeneity in equation (2) allows us to estimate the effects of Blueprint interventions largely free from any negative selection bias that may be affecting our results in Tables 10 and 11.

We first show the effects of the total number of Blueprint interventions on the number of required courses completed by each student in each year. We present separate estimates for summer interventions and school-year interventions. Because the effects of Blueprint interventions are likely to be most severe when they take away class time during the school year, the latter estimate may be of the most interest. The first column in Table 12 details our results. The second column in Table 12 breaks down the Blueprint interventions into the four main high school interventions: Literacy Block, Literacy Core, Literacy Block/Core for EL students, and Blueprint Summer school.

[8] With the student-fixed-effects model, we no longer include controls for students' grade 8 test scores. This allows us to include students into our fixed effects analysis who were excluded from the previous models because of missing test-score data.

Table 12 shows that as one might expect, the Summer school element bears no relation to the number of A–G courses taken during the school year. Conversely, Blueprint interventions during the school year have significant effects on student class-taking behavior within years. These interventions are associated with a roughly one-for-one reduction in the number of required courses taken by students in the year of the intervention. However, a comparison of the results from Table 12 with those from Table 11 indicates that the within-year reductions in the number of courses taken by Blueprint participants do not translate into a reduction in the number of participants who ultimately complete these requirements. That is, Table 11 shows that conditional on Blueprint participation, additional Blueprint interventions do not affect whether students complete the course requirements despite the short-term displacement effects.

If students are not close to completing the A–G sequence in general, this may offer one explanation for why Blueprint interventions have a limited effect on overall completion rates despite their strong within-year effects. Table 13 provides summary statistics for Blueprint participants in terms of their progress in completing the course requirements. The table shows that almost three-quarters of the students participating in the Blueprint interventions end up three or more subject requirements away from completing the A–G sequence. Because each of the subject requirements involves the passage of two to eight high-school semester courses, we infer that most of these students would have fallen far short of completing the admission requirements with or without the Blueprint.

Finally, we examine the nature of the class-substitution behavior of Blueprint participants within years (Table 19). We find that students who participate in Summer school Blueprint interventions are actually more likely to take most A–G classes (the strong exceptions being classes for the foreign language requirement and to a lesser degree, science). For students who are forced to substitute out of classes due to school-year Blueprint interventions, the most common substitution is out of foreign-language classes. This result is quite intuitive if these students are already struggling with English. Students also tend to substitute out of art classes but are more likely to complete university-required courses in English and, to some degree, in math, and in social studies. This implies that, as indicated by Table 13, Blueprint interventions encourage participants to take additional classes in the two subjects in which Blueprint students are least likely to complete their university course requirements, English and math. In some ways, because these are core subjects, this is a salutary outcome.

Conclusion

Additional time on task for elementary and middle school students in general boosted reading achievement. The most impressive effect was that Literacy Core for non-EL students in middle school was associated with a gain of 0.16 standard deviations in reading in the year the student participated. This is a sizeable gain. The findings validate the idea that extra time on task for students who are behind in reading can lead to meaningful gains in literacy.

However, similar programs at the high school level appear to have slowed down gains in reading achievement for participants. It is impossible to know for sure why the programs were so much less effective in the higher grades—but we offer some ideas in the main report.

Policymakers often design interventions to narrow the achievement gap in the hope and belief that students with less experienced teachers might benefit the most. As mentioned earlier, such a pattern may or may not obtain depending on whether the intervention is a substitute for—or a complement to—highly experienced teachers. We found abundant evidence that elementary school students with relatively inexperienced teachers gain less from the interventions. In particular, this applied for the peer coach element, suggesting that in elementary schools, the coaches might have been most complementary to relatively seasoned homeroom teachers. At the middle school and high school level, the hypothesis of no interactions was handily retained, although in a few middle school cases we again found some evidence that more experienced teachers were associated with larger effects of some of the interventions.

An important aspect of the reform was teacher professional development, operating largely through the placement of peer coaches in each school to help teach teachers about pedagogical tools for boosting literacy. Our simplest models found no effects of variation in the ratio of peer coaches as a percentage of enrollment at each school, but more complex models did suggest that peer coaching tended to become a positive and significant contributor to students' reading gains in the later years of the program. In addition, at the elementary school level, we found that the overall effect of peer coaches was estimated to be positive once we took into account the possibility that the effect of peer coaching could vary with the experience of the homeroom teacher. As related above, elementary students with more experienced teachers were more likely to gain from increases in the intensity of peer coaching at their schools.

In sum, the Blueprint reforms successfully boosted reading achievement in elementary and middle schools, but not high schools, and did not appreciably hurt student performance or engagement in other subject areas.

TABLE 1 Student, family, and neighborhood controls used in the statistical models for elementary students

Student characteristics
Fixed effects for each student to control for all characteristics of a student that are fixed over time, such as race. Controls for the student's standardized reading score in the given subject last year; controls for students who changed schools that year, or switched schools unexpectedly; age; grade level, year dummies.
Family characteristics
Controls for the level of education of the more highly educated parent.
Neighborhood characteristics
Fixed effects for student's home zip code.

TABLE 2 School, classroom, and student body controls used in the statistical models for elementary students

School characteristics
Fixed effects for each school to control for all fixed characteristics of the school. Controls for whether the school was a year-round school.
Student body characteristics at the school level
% eligible for free/reduced-price meal; separate controls for % of students who are Hispanic, black, Asian, Pacific Islander, native American; % of students who are EL, % Fluent English Proficient (FEP); controls for student mobility: % who changed schools that year, who switched schools unexpectedly, and who were new to the district.
Student body characteristics at the grade level
Mean test scores in previous spring's test of all students in the student's current grade, standardized to district average.
Classroom and teacher characteristics
Class size; controls for teacher characteristics: interactions of credentials (intern, emergency credential, full credential) with indicators of years of teaching experience (e.g. 0-2, 3-5, 6-9); master's degree, Ph.D.; bachelor's in math, English, social science, science, language, other major (except education) (separate variables for each major); the Crosscultural Language and Academic Development credential (CLAD), (Spanish) Bilingual CLAD (BCLAD), CLAD alternative credential, BCLAD alternative credential; controls for teachers who are black, Asian, Hispanic, other non-white, and female.

TABLE 3 Estimates of student participation in Blueprint programs (%) and the extent of the peer coach program by grade and year

EDRP

Grade level	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
1	31.1	34.8	29.1	5.8	5.0
2	33.0	32.7	25.4	5.4	3.7
3	41.2	34.2	23.4	6.0	3.6
4	27.7	29.6	21.8	6.2	3.3
5	22.4	26.4	18.0	4.9	3.1
6	15.6	18.2	14.1	3.5	3.9
7	9.5	11.7	9.9	5.5	5.1
8	9.8	10.0	9.9	4.3	4.9
9	7.4	0.8	0.9	0	0
10	1.6	0.1	0	0	0

Genre Studies

Grade level	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
6	0.1	31.1	36.9	42.0	37.8	41.6
7	0	0	0	6.5	6.8	12.9
8	0	0	0	6.2	4.1	12.9
9	0	0	0	0	0	0
10	0	0	0	0	0	0
11	0	0	0	0	0	0
12	0	0	0	0	0	0

Summer school

Grade level	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
1	17.6	31.2	33.0	38.2	20.6
2	15.6	28.6	30.4	36.1	19.8
3	16.5	26.8	29.2	33.6	20.1
4	8.6	23.8	26.8	33.3	18.6
5	14.5	23.0	25.7	30.1	15.2
6	20.8	24.2	25.7	24.4	11.1
7	4.1	22.1	27.3	21.4	10.0
8	6.8	20.9	30.5	29.4	16.9
9	14.2	22.8	20.7	17.6	5.7
10	0	10.5	17.9	10.6	0.7
11	0	2.0	3.5	2.5	0.2
12	0	0.4	1.1	0.8	0.1

Intersession

Grade level	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
1	17.6	18.7	19.3	26.1	0
2	15.7	16.0	17.2	23.6	13.3
3	16.5	16.0	17.4	22.8	12.1
4	8.6	13.7	14.6	22.9	12.4
5	14.5	14.4	13.7	20.7	11.1
6	16.7	14.7	15.2	15.7	8.7
7	6.7	5.2	6.8	6.7	4.9
8	6.6	6.0	6.4	6.9	0

Continued on next page

Literacy Block

Grade level	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
6	0	13.3	8.5	7.9	10.1	9.5
7	0	22.2	19.1	17.8	17.7	12.6
8	0	17.5	16.0	17.8	18.6	12.9
9	23.1	12.8	12.7	12.8	15.8	15.9
10	3.9	20.3	20.1	19.5	17.3	10.1
11	0	3.0	1.4	1.7	0.3	0.1
12	0	0.1	0	0	0	0.1

Literacy Core

Grade level	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
6	0	0.6	1.2	0.9	0
7	0	0.5	0.4	0.2	0
8	0	0.1	0.1	0	0
9	9.0	7.2	6.8	0.1	0
10	0.2	0.2	0.2	0.1	0
11	0	0	0	0	0
12	0	0	0	0	0

Literacy Block and Literacy Core for EL students

Grade level	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
6	0	10.5	10.9	10.2	13.3	11.8
7	0	20.5	18.4	14.3	16.2	15.4
8	0	17.0	14.0	11.2	13.5	13.3
9	11.1	15.9	14.5	12.9	13.8	14.4
10	5.3	12.7	14.9	12.9	11.9	9.2
11	2.1	4.9	3.4	3.5	2.6	2.1
12	0.7	0.7	0.7	1.0	0.7	0.5

Mean peer coach as percent of enrollment

Grade level	1999-2000	2000-2001	2001-2002	2002-2003	2003-2004	2004-2005
1	0.10	0.31	0.19	0.20	0.17	0.18
2	0.10	0.31	0.19	0.20	0.17	0.18
3	0.09	0.31	0.19	0.20	0.17	0.19
4	0.09	0.30	0.19	0.20	0.17	0.18
5	0.09	0.30	0.18	0.20	0.17	0.17
6	0.01	0.19	0.12	0.14	0.11	0.12
7	0.07	0.12	0.08	0.10	0.10	0.08
8	0.07	0.12	0.08	0.10	0.10	0.08
9	0.07	0.09	0.05	0.05	0.03	0.01
10	0.07	0.09	0.05	0.05	0.02	0.01
11	0.07	0.08	0.05	0.05	0.02	0.01
12	0.07	0.09	0.05	0.05	0.02	0.01

NOTES: Unless specifically listed, participation in all programs was 0 in 1999–2000. The tables for Literacy Block and Literacy Core show the percentage of all students who were enrolled in Literacy Block and Core courses intended for non–English Learners (non-EL), while the table for Literacy Block/Core for EL shows the percentage of all students who were EL students enrolled in Block or Core. Thus, to estimate the total percentage of all students in a given grade and year who were in Block and Core, one should add together the cells in those three tables. The final table shows the student-based average of the number of peer coaches divided by enrollment at the school, expressed as a percentage.

TABLE 4 Pooled models for elementary students using gains in reading scores

	Model i	Model ii	Model iii	Model iv	Model v	Model vi
Class size	No	Yes	No	Yes	Yes	Yes
Teacher qualifications	No	No	Yes	Yes	Yes	Yes
Interactions1:	No	No	No	No	Yes	Yes
Block, Core, peer coach						
Interactions2:	No	No	No	No	No	Yes
EDRP, summer						
Average years experience of peer coach	-0.0002	-0.0002	-0.0002	-0.0002	-0.0001	-0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
API-2 elementary school	0.0325	0.0339	0.0301	0.0315	0.0496	0.0471
	(0.0073)**	(0.0073)**	(0.0073)**	(0.0073)**	(0.0107)**	(0.0107)**
Focus school	0.0231	0.0242	0.0222	0.0233	0.0129	0.0116
	(0.0089)**	(0.0089)**	(0.0089)*	(0.0089)**	(0.0123)	(0.0123)
Peer coach as % of enrollment	0.0301	0.0299	0.0294	0.0293	0.0401	0.0390
	(0.0167)	(0.0167)	(0.0167)	(0.0167)	(0.0168)*	(0.0169)*
EDRP participation	-0.0025	-0.0029	-0.0017	-0.0021	-0.0015	0.0082
	(0.0047)	(0.0047)	(0.0047)	(0.0047)	(0.0047)	(0.0071)
Summer school participation	0.0023	0.0024	0.0007	0.0007	0.0012	0.0064
	(0.0072)	(0.0072)	(0.0072)	(0.0072)	(0.0072)	(0.0083)
Intersession participation	0.0312	0.0312	0.0307	0.0307	0.0297	0.0320
	(0.0091)**	(0.0090)**	(0.0090)**	(0.0090)**	(0.0090)**	(0.0091)**
Interactions of Blueprint variables with teacher characteristics						
Peer coach and 0-2 years experience					-0.0004	-0.0003
					(0.0004)	(0.0004)
Peer coach and 3-5 years experience					-0.0016	-0.0016
					(0.0004)**	(0.0004)**
Peer coach and 6-9 years experience					-0.0007	-0.0007
					(0.0004)	(0.0004)
Focus school and 0-2 years experience					0.0001	0.0001
					(0.0002)	(0.0002)
Focus school and 3-5 years experience					0.0004	0.0004
					(0.0001)**	(0.0001)**
Focus school and 6-9 years experience					-0.0000	-0.0000
					(0.0002)	(0.0002)

Continued on next page

Continued

API-2 school and 0-2 years experience					-0.0002	-0.0002
					(0.0001)	(0.0001)
API-2 school and 3-5 years experience					-0.0003	-0.0003
					(0.0001)*	(0.0001)*
API-2 school and 6-9 years experience					-0.0001	-0.0000
					(0.0001)	(0.0001)
EDRP and 0-2 years experience						-0.0003
						(0.0001)*
EDRP and 3-5 years experience						-0.0001
						(0.0001)
EDRP and 6-9 years experience						-0.0001
						(0.0001)
Summer session and 0-2 years experience interaction						-0.0002
						(0.0001)*
Summer session and 3-5 years experience interaction						-0.0001
						(0.0001)
Summer session and 6-9 years experience interaction						0.0000
						(0.0001)
Observations	141220	141220	141220	141220	141220	141220
Number of student id's	74456	74456	74456	74456	74456	74456
R-squared	0.66	0.66	0.66	0.66	0.66	0.66
P-value for F-test for no student fixed effects	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
P-value for F-test for exclusion of Blueprint experience interactions					<0.0001	<0.0001

NOTES: Each column represents reading scores, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 5 Pooled models for middle school students using gains in reading scores

	Model i	Model ii	Model iii	Model iv	Model v	Model vi
Class size	No	Yes	No	Yes	Yes	Yes
Teacher qualifications	No	No	Yes	Yes	Yes	Yes
Interactions1:	No	No	No	No	Yes	Yes
Block, Core, peer coach						
Interactions2:	No	No	No	No	No	Yes
EDRP, summer						
Literacy Block	0.0511	0.0460	0.0566	0.0501	0.0506	0.0491
	(0.0061)**	(0.0063)**	(0.0062)**	(0.0065)**	(0.0090)**	(0.0090)**
Literacy Core	0.1607	0.1535	0.1713	0.1642	0.1280	0.1275
	(0.0412)**	(0.0413)**	(0.0412)**	(0.0413)**	(0.0578)*	(0.0578)*
Literacy Block/Core for EL students	0.0379	0.0339	0.0474	0.0406	0.0404	0.0397
	(0.0080)**	(0.0082)**	(0.0081)**	(0.0083)**	(0.0083)**	(0.0083)**
Genre Studies	0.0710	0.0696	0.0687	0.0653	0.0643	0.0640
	(0.0061)**	(0.0061)**	(0.0062)**	(0.0063)**	(0.0068)**	(0.0068)**
EDRP	0.0015	0.0013	0.0015	0.0014	0.0012	0.0143
	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0079)
Summer session	0.0058	0.0054	0.0053	0.0053	0.0053	0.0027
	(0.0055)	(0.0055)	(0.0055)	(0.0055)	(0.0055)	(0.0063)
Intersession	0.0091	0.0106	0.0077	0.0087	0.0090	0.0089
	(0.0092)	(0.0092)	(0.0092)	(0.0092)	(0.0093)	(0.0094)
Average years experience of peer coach	-0.0001	-0.0001	0.0001	0.0001	0.0001	0.0001
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Peer coach as % of enrollment	0.0320	0.0328	0.0247	0.0220	0.0238	0.0227
	(0.0278)	(0.0278)	(0.0279)	(0.0279)	(0.0287)	(0.0288)
Accelerated class	-0.0430	-0.0437	-0.0499	-0.0510	-0.0489	-0.0489
	(0.0255)	(0.0255)	(0.0255)	(0.0255)*	(0.0255)	(0.0256)
Interactions of Blueprint variables with teacher characteristics						
Genre Studies and 0-2 years teacher experience					-0.0007	-0.0007
					(0.0003)**	(0.0003)**
Genre Studies and 3-5 years teacher experience					0.0001	0.0001
					(0.0001)	(0.0001)
Genre Studies and 6-9 years teacher experience					0.0001	0.0001
					(0.0001)	(0.0001)

Continued on next page

Continued

Literacy Block and 0-2 years teacher experience					0.0000	0.0000
					(0.0001)	(0.0001)
Literacy Block and 3-5 years teacher experience					-0.0000	-0.0000
					(0.0001)	(0.0001)
Literacy Block and 6-9 years teacher experience					-0.0000	0.0000
					(0.0001)	(0.0001)
Literacy Core and 0-2 years teacher experience					0.0004	0.0005
					(0.0010)	(0.0010)
Literacy Core and 3-5 years teacher experience					-0.0003	-0.0004
					(0.0010)	(0.0010)
Literacy Core and 6-9 years teacher experience					0.0015	0.0015
					(0.0009)	(0.0009)
Peer coach and 0-2 years teacher experience					0.0003	0.0003
					(0.0008)	(0.0008)
Peer coach and 3-5 years teacher experience					0.0001	0.0001
					(0.0008)	(0.0008)
Peer coach and 6-9 years teacher experience					-0.0002	-0.0001
					(0.0008)	(0.0008)
EDRP and 0-2 years teacher experience						-0.0004
						(0.0002)*
EDRP and 3-5 years teacher experience						-0.0004
						(0.0002)**
EDRP and 6-9 years teacher experience						-0.0001
						(0.0001)
Summer school and 0-2 years teacher experience						0.0000
						(0.0001)

Continued on next page

Continued

Summer school and 3-5 years teacher experience						0.0001
						(0.0001)
Summer school and 6-9 years teacher experience						0.0000
						(0.0001)
Observations	115280	115280	115280	115280	115280	115280
Number of student id's	61619	61619	61619	61619	61619	61619
R-squared	0.64	0.64	0.64	0.64	0.64	0.64
P-value for F-test for no student fixed effects	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
P-value for F-test for exclusion of Blueprint-experience interactions					0.3620	0.1678

NOTES: Each column represents reading scores, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 6 Pooled models for high school students using gains in reading scores

	Model i	Model ii	Model iii	Model iv	Model v	Model vi
Class size	No	Yes	No	Yes	Yes	Yes
Teacher qualifications	No	No	Yes	Yes	Yes	Yes
Interactions1:	No	No	No	No	Yes	Yes
Block, Core, peer coach						
Interactions2:	No	No	No	No	No	Yes
EDRP, summer						
Literacy Block	-0.0435	-0.0427	-0.0407	-0.0406	-0.0452	-0.0454
	(0.0084)**	(0.0085)**	(0.0085)**	(0.0085)**	(0.0116)**	(0.0116)**
Literacy Core	-0.0979	-0.0971	-0.0982	-0.0979	-0.1290	-0.1294
	(0.0193)**	(0.0194)**	(0.0194)**	(0.0194)**	(0.0324)**	(0.0325)**
Literacy Block/Core for EL students	-0.1731	-0.1688	-0.1685	-0.1651	-0.1650	-0.1649
	(0.0117)**	(0.0118)**	(0.0118)**	(0.0119)**	(0.0119)**	(0.0119)**
Summer session	-0.0233	-0.0219	-0.0227	-0.0214	-0.0205	-0.0194
	(0.0077)**	(0.0077)**	(0.0077)**	(0.0077)**	(0.0078)**	(0.0093)*
Average years experience of peer coach	-0.0009	-0.0009	-0.0009	-0.0009	-0.0009	-0.0009
	(0.0003)**	(0.0003)**	(0.0003)**	(0.0003)**	(0.0003)**	(0.0003)**
Peer coach as % of enrollment	0.0700	0.0905	0.0536	0.0662	0.0774	0.0773
	(0.0534)	(0.0536)	(0.0538)	(0.0540)	(0.0557)	(0.0557)
Interactions of Blueprint variables with teacher characteristics						
Literacy Block and 0-2 years teacher experience					0.0001	0.0001
					(0.0002)	(0.0002)
Literacy Block and 3-5 years teacher experience					-0.0000	-0.0000
					(0.0002)	(0.0002)
Literacy Block and 6-9 years teacher experience					0.0001	0.0001
					(0.0002)	(0.0002)
Literacy Core and 0-2 years teacher experience					0.0002	0.0003
					(0.0004)	(0.0004)
Literacy Core and 3-5 years teacher experience					0.0008	0.0008
					(0.0005)	(0.0005)

Continued on next page

Continued

Literacy Core and 6-9 years teacher experience					0.0005	0.0005
					(0.0006)	(0.0006)
Peer coach and 0-2 years teacher experience					-0.0009	-0.0009
					(0.0012)	(0.0012)
Peer coach and 3-5 years teacher experience					-0.0024	-0.0023
					(0.0015)	(0.0015)
Peer coach and 6-9 years teacher experience					0.0003	0.0003
					(0.0015)	(0.0015)
Summer school and 0-2 years teacher experience						-0.0002
						(0.0003)
Summer school and 3-5 years teacher experience						0.0000
						(0.0002)
Summer school and 6-9 years teacher experience						-0.0000
						(0.0002)
Observations	97656	97656	97656	97656	97656	97656
Number of student id's	53416	53416	53416	53416	53416	53416
R-squared	0.61	0.61	0.61	0.61	0.61	0.61
P-value for F-test for no student fixed effects	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
P-value for F-test for exclusion of Blueprint-experience interactions					0.6474	0.8252

NOTES: Each column represents reading scores, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 7 Math achievement outcomes for all students by grade span using pooled models of specification 4

	Elementary	Middle	High
Class size	Yes	Yes	Yes
Teacher qualifications	Yes	Yes	Yes
Interactions1:	No	No	No
Block, Core, peer coach			
Interactions2:	No	No	No
EDRP, summer			
Average years experience of peer coach	0.0009	-0.0018	0.0000
	(0.0003)**	(0.0003)**	(0.0004)
API-2 elementary school	0.0023		
	(0.0088)		
Focus school	0.0161		
	(0.0107)		
Peer coach as % of enrollment	0.0262	0.0724	0.0247
	(0.0200)	(0.0410)	(0.0962)
EDRP participation	0.0239	-0.0225	
	(0.0056)**	(0.0082)**	
Summer school participation	0.0051	0.0287	-0.0128
	(0.0086)	(0.0072)**	(0.0102)
Intersession participation	0.0802	-0.0284	
	(0.0109)**	(0.0127)*	
Literacy Block		0.0160	-0.0433
		(0.0086)	(0.0114)**
Literacy Core		0.0088	-0.0576
		(0.0510)	(0.0250)*
Block/Core for EL students		-0.0255	-0.0697
		(0.0116)*	(0.0165)**
Genre Studies		0.0184	
		(0.0086)*	
Observations	145716	96429	77852
Number of student id's	75818	53032	43986
R-squared	0.65	0.61	0.71
P-value for F-test for no student fixed effects	<0.0001	<0.0001	<0.0001
Mean of dependent variable	0.0278	-0.0268	-0.0554
Standard deviation of dependent variable	0.6827	0.6230	0.7766

NOTES: Each column represents math scores, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 8 Pooled models using percent of total school year absent as dependent variable

	Elementary	Middle	High
Average years experience of peer coach	-0.0007	-0.0083	0.0111
	(0.0010)	(0.0018)**	(0.0015)**
API-2 Elementary school	-0.0825		
	(0.0364)*		
Focus school	-0.0456		
	(0.0408)		
Peer coach as % of enrollment	0.4189	0.4343	-1.2229
	(0.0774)**	(0.2209)*	(0.3220)**
EDRP participation	-0.1333	-0.1351	
	(0.0213)**	(0.0496)**	
Summer school participation	-0.0549	0.1323	-0.0973
	(0.0353)	(0.0447)**	(0.0549)
Intersession participation	-0.2308	-0.2806	
	(0.0439)**	(0.0731)**	
Literacy Block		0.2861	0.0364
		(0.0531)**	(0.0583)
Literacy Core		0.2684	-0.4113
		(0.2693)	(0.1279)**
Block/Core for EL students		0.3829	-0.3596
		(0.0651)**	(0.0724)**
Genre Studies		0.3422	
		(0.0517)**	
Observations	267963	148263	164285
Number of student id's	118795	78421	77289
R-squared	0.02	0.06	0.07
Mean of dependent variable:	4.42	5.39	4.83
Standard deviation of dependent variable	4.38	6.32	6.04

NOTES: Each column represents percent of time absent, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 9 Student fixed-effect linear probability models of being retained

	Elementary	Middle	High
API-2 elementary school	-0.0014		
	(0.0014)		
Focus school	0.0001		
	(0.0016)		
EDRP participation	0.0019	-0.0004	
	(0.0008)*	(0.0012)	
Summer school participation	-0.0105	0.0024	-0.0103
	(0.0013)**	(0.0012)*	(0.0027)**
Intersession participation	0.0078	-0.0000	
	(0.0017)**	(0.0019)	
Literacy Block		0.0007	-0.0072
		(0.0013)	(0.0028)**
Literacy Core		-0.0206	-0.0363
		(0.0070)**	(0.0059)**
Block/Core for EL students		0.0154	-0.0200
		(0.0017)**	(0.0036)**
Genre Studies		0.0014	
		(0.0015)	
Observations	244218	123904	131959
Number of student id's	97940	67522	64202
R-squared	0.06	0.50	0.23

NOTES: Each column represents grade retention, using school and student fixed effects models. Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 10 Dependent variable: indicator for graduation

Blueprint participation indicator	-0.087 (0.005)**	0.017 (0.006)*	-0.022 (0.006)**	-0.132 (0.011)**	-0.060 (0.010)**	-0.057 (0.010)**
Number of additional Blueprint interventions				0.026 (0.005)**	0.045 (0.005)**	0.021 (0.005)**
Student-level covariates		X	X		X	X
School-level covariates			X			X
R^2	0.02	0.15	0.22	0.02	0.16	0.22
Observations	21,192	21,025	20,777	21,192	21,025	20,777

NOTE: The mean of the dependent variable, an indicator for high-school graduation, was 0.89 for the full sample.

** Denotes significance at the 1 percent level of confidence.

* Denotes significance at the 5 percent level of confidence.

TABLE 11 Dependent variable: indicator for the completion of A–G classes

Blueprint participation indicator	-0.272 (0.006)**	-0.048 (0.006)**	-0.078 (0.007)**	-0.238 (0.011)**	-0.068 (0.010)**	-0.072 (0.011)**
Number of additional Blueprint interventions				-0.019 (0.005)**	0.012 (0.005)*	-0.004 (0.005)
Student-level covariates		X	X		X	X
School-level covariates			X			X
R^2	0.07	0.28	0.31	0.07	0.28	0.31
Observations	21,192	21,025	20,777	21,192	21,025	20,777

NOTE: The mean of the dependent variable, an indicator for completion of the A–G course sequence, was 0.36 for the full sample.

** Denotes significance at the 1 percent level of confidence.

* Denotes significance at the 5 percent level of confidence.

TABLE 12 Results from equation (2). Dependent variable: number of A–G classes taken in a given year

Overall		
Total school-year Blueprint interventions	-0.96 (0.03)**	
Summer Blueprint interventions	-0.05 (0.04)	
By intervention		
Total Block interventions		-0.90 (0.04)**
Total Core Interventions		-1.02 (0.08)**
Total Block/Core interventions for EL students		-1.07 (0.06)**
Total summer interventions		-0.05 (0.04)
R^2	0.15	0.15
Observations	140,037	140,037

NOTE: See the text for a full list of regressors.

**Denotes significance at the 1 percent level of confidence.

* Denotes significance at the 5 percent level of confidence.

TABLE 13 Summary statistics for Blueprint participants

Progress on university course requirements	Share of Blueprint participants
Completed all	13.7%
Completed 5 / 7	12.7%
Completed 4 / 7	12.4%
Completed 3 / 7	14.6%
Completed 2 / 7	16.2%
Completed one or less	30.5%
Progress by requirement	
Completed Math	32.4%
Completed English	35.0%
Completed Science	45.2%
Completed Social Studies	72.8%
Completed Art	57.5%
Completed Foreign Language	39.0%

NOTE: The seventh university course requirement requires that students take two additional semesters in any university eligible classes in any of the six required subjects. We assume that no student fulfills this final requirement until the first six requirements are complete.

TABLE 14 Summary statistics for models in Tables 4, 5, and 6

	Elementary		Middle		High	
Variable	Mean	SD	Mean	SD	Mean	SD
Gain in reading test score	0.0468	0.5607	0.0126	0.5042	-0.0158	0.5371
Lagged test score	0.0897	0.9812	0.1484	0.9791	0.2078	0.9590
Average years experience of peer coach	14.3013	10.0177	12.4868	9.9454	13.860	12.021
API 2 elementary school	0.1050	0.3066				
Focus school	0.0708	0.2565				
Peer coach as % of enrollment	0.1903	0.1330	0.1024	0.0823	0.0474	0.0528
EDRP participation	0.1525	0.3595	0.0782	0.2684		
Summer school participation	0.1946	0.3959	0.1681	0.3740	0.0796	0.2707
Intersession participation	0.1242	0.3300	0.0660	0.2483		
Literacy Block			0.1103	0.3069	0.0988	0.2912
Literacy Core			0.0015	0.0380	0.0106	0.1012
Block/Core for EL students			0.0978	0.2922	0.0688	0.2481
Genre Studies			0.1530	0.3552		

TABLE 15 Pooled models by grade span with and without lag test score

	Elementary	Elementary No lag score	Middle	Middle No lag score	High	High No lag score
Class size	Yes	Yes	Yes	Yes	Yes	Yes
Teacher qualifications	Yes	Yes	Yes	Yes	Yes	Yes
Interactions1:	No	No	No	No	No	No
Block, Core, peer coach						
Interactions2:	No	No	No	No	No	No
EDRP, summer						
Average years experience of peer coach	-0.0002	0.0007	0.0002	0.0009	-0.0009	-0.0018
	(0.0002)	(0.0004)*	(0.0003)	(0.0004)*	(0.0003)**	(0.0004)**
Peer coach as % of enrollment	0.0293	0.0153	0.0220	0.0628	0.0662	0.0203
	(0.0167)	(0.0281)	(0.0415)	(0.0459)	(0.0540)	(0.0850)
Summer school participation	0.0007	0.0191	-0.0011	0.0507	-0.0214	-0.0327
	(0.0072)	(0.0121)	(0.0065)	(0.0091)**	(0.0077)**	(0.0122)**
API-2 elementary school	0.0315	0.0356				
	(0.0073)**	(0.0124)**				
Focus school	0.0233	0.0673				
	(0.0089)**	(0.0150)**				
EDRP participation	-0.0021	0.1290	0.0001	0.0583		
	(0.0047)	(0.0079)**	(0.0077)	(0.0100)**		
Intersession participation	0.0307	0.0879	0.0147	0.0115		
	(0.0090)**	(0.0152)**	(0.0125)	(0.0152)		
Literacy Block			0.0570	0.0744	-0.0406	0.0768
			(0.0078)**	(0.0106)**	(0.0085)**	(0.0134)**
Literacy Core			0.1600	0.3195	-0.0979	0.0488
			(0.0433)**	(0.0679)**	(0.0194)**	(0.0305)
Literacy Block/Core for EL students			0.0058	0.1120	-0.1651	0.0678
			(0.0114)	(0.0136)**	(0.0119)**	(0.0186)**
Genre Studies			0.0601	0.0070		
			(0.0085)**	(0.0103)		
Observations	141220	141220	115280	115280	97656	97656
Number of student id's	74456	74456	61619	61619	53416	53416
R-squared	0.09	0.03	0.07	0.03	0.03	0.03
P-value for F-test for no student fixed effects	<0.0001	1.0	<0.0001	1.0	<0.0001	1.0

NOTES: Each column represents reading scores, using school and student fixed effects models.

Standard errors in parentheses.

* Significant at 5%.

** Significant at 1%.

TABLE 16 Pooled models with year variations for elementary students

	Model iv	Tests that overall effect equals 0 in given year
Class size	Yes	
Teacher qualifications	Yes	
Interactions1:	No	
Block, Core, peer coach		
Interactions2:	No	
EDRP, summer		
Focus school	-0.0007	
	(0.0143)	
Focus and 2001-2002 school year	0.0286	*
	(0.0153)	
Focus and 2002-2003 school year	0.1191	**
	(0.0199)**	
Focus and 2003-2004 school year	0.0262	
	(0.0200)	
Focus and 2004-2005 school year	0.0076	
	(0.0273)	
API-2 elementary school	0.0034	
	(0.0119)	
API-2 and 2001-2002 school year	0.0670	**
	(0.0147)**	
API-2 and 2002-2003 school year	0.0644	**
	(0.0145)**	
API-2 and 2003-2004 school year	0.0183	
	(0.0166)	
API-2 and 2004-2005 school year	-0.0061	
	(0.0205)	
EDRP participation	-0.0144	
	(0.0080)	
EDRP and 2001-2002 school year	0.0218	
	(0.0109)*	
EDRP and 2002-2003 school year	0.0191	
	(0.0116)	
EDRP and 2003-2004 school year	0.0037	
	(0.0179)	
EDRP and 2004-2005 school year	0.0554	
	(0.0261)*	
Summer school participation	0.3496	
	(0.3081)	
Summer and 2001-2002 school year	-0.3480	
	(0.3084)	
Summer and 2002-2003 school year	-0.3612	
	(0.3083)	
Summer and 2003-2004 school year	-0.3559	
	(0.3084)	
Summer and 2004-2005 school year	-0.2934	**
	(0.3086)	

Continued on next page

Continued

Intersession participation	-0.3081	
	(0.3083)	
Intersession and 2001-2002 school year	0.3279	
	(0.3086)	
Intersession and 2002-2003 school year	0.3501	**
	(0.3085)	
Intersession and 2003-2004 school year	0.3321	
	(0.3086)	
Intersession and 2004-2005 school year	0.3264	
	(0.3091)	
Average years experience of peer coach	-0.0002	
	(0.0002)	
Peer coach as % of enrollment	0.0525	*
	(0.0229)*	
Peer coach and 1999-2000 school year	-0.1273	
	(0.0482)**	
Peer coach and 2001-2002 school year	-0.1279	
	(0.0479)**	
Peer coach and 2002-2003 school year	0.0507	**
	(0.0438)	
Peer coach and 2003-2004 school year	-0.0006	
	(0.0416)	
Peer coach and 2004-2005 school year	0.1031	**
	(0.0473)*	
Observations	141220	
Number of student id's	74456	
R-squared	0.66	

P-values for F-tests that selected year interactions equal zero:

All Blueprint variables	<0.0001
Focus schools	<0.0001
API-2 schools	<0.0001
EDRP	0.0916
Summer	0.0032
Intersession	0.5321
Peer coach as % of enrollment	0.0001

NOTES: Standard errors in parentheses. The rightmost column provides a test that the overall effect of the given Blueprint element was zero in the given year. Rows that do not indicate a year refer to a test that the given Blueprint element had a zero effect in the base year, which was 2000-2001 for all Blueprint elements except for peer coaches, for which the base year was 1999-2000.

* Significant at 5%.

** Significant at 1%.

TABLE 17 Pooled models with year variations for middle school students

	Model iv	Tests that overall effect equals 0 in a given year
Class size	Yes	
Teacher qualifications	Yes	
Interactions1:	No	
Block, Core, peer coach		
Interactions2:	No	
EDRP, summer		
Literacy Block	0.0457	**
	(0.0102)**	
Literacy Block and 2001-2002 school year	-0.0140	*
	(0.0136)	
Literacy Block and 2002-2003 school year	-0.0130	*
	(0.0140)	
Literacy Block and 2003-2004 school year	-0.0050	*
	(0.0146)	
Literacy Block and 2004-2005 school year	0.0009	**
	(0.0179)	
Literacy Core	0.1450	**
	(0.0571)*	
Literacy Core and 2002-2003 school year	-0.1022	
	(0.0747)	
Literacy Core and 2003-2004 school year	0.0534	**
	(0.0821)	
Literacy Block/Core for EL students	-0.0134	
	(0.0116)	
Literacy Block/Core for EL students and 2001-2002 school year	0.0119	
	(0.0141)	
Literacy Block/Core for EL students and 2002-2003 school year	0.0753	**
	(0.0161)**	
Literacy Block/Core for EL students and 2003-2004 school year	0.0880	**
	(0.0173)**	
Literacy Block/Core for EL students and 2004-2005 school year	0.1643	**
	(0.0196)**	
Genre Studies	0.0863	**
	(0.0123)**	
Genre Studies and 1999-2000 school year	-0.0207	
	(0.0364)	
Genre Studies and 2001-2002 school year	0.0245	**
	(0.0155)	
Genre Studies and 2002-2003 school year	-0.0251	**
	(0.0157)	
Genre Studies and 2003-2004 school year	-0.0486	*
	(0.0149)**	

Continued on next page

Continued

Genre Studies and 2004-2005 school year	-0.0710	
	(0.0185)**	
EDRP	-0.0068	
	(0.0115)	
EDRP and 2001-2002 school year	0.0026	
	(0.0153)	
EDRP and 2002-2003 school year	0.0069	
	(0.0158)	
EDRP and 2003-2004 school year	0.0301	
	(0.0208)	
EDRP and 2004-2005 school year	0.0532	*
	(0.0245)*	
Summer session	-0.0157	
	(0.0155)	
Summer and 2001-2002 school year	0.0086	
	(0.0181)	
Summer and 2002-2003 school year	0.0232	
	(0.0178)	
Summer and 2003-2004 school year	0.0119	
	(0.0187)	
Summer and 2004-2005 school year	0.0332	
	(0.0202)	
Intersession	0.0039	
	(0.0167)	
Intersession and 2001-2002 school year	0.0410	**
	(0.0209)*	
Intersession and 2002-2003 school year	0.0338	*
	(0.0212)	
Intersession and 2003-2004 school year	-0.0106	
	(0.0232)	
Intersession and 2004-2005 school year	0.0000	
	(0.0000)	
Average years experience of peer coach	0.0003	
	(0.0002)	
Peer coach as % of enrollment	-0.0269	
	(0.0503)	
Peer coach and 1999-2000 school year	-0.3525	**
	(0.0831)**	
Peer coach and 2001-2002 school year	0.1920	**
	(0.0775)*	
Peer coach and 2002-2003 school year	0.1601	*
	(0.0703)*	
Peer coach and 2003-2004 school year	0.0538	
	(0.0713)	
Peer coach and 2004-2005 school year	0.0813	
	(0.0918)	
Observations	115280	
Number of stud id	61619	
R-squared	0.64	

Continued on next page

Continued

P-values for F-tests that selected year interactions equal zero:	
All Blueprint variables	<0.0001
Literacy Block	0.8072
Literacy Core	0.1598
Literacy Block/Core for EL	<0.0001
Genre Studies	<0.0001
EDRP	0.1706
Summer	0.3651
Intersession	0.0360
Peer coach as % of enrollment	<0.0001

NOTES: Standard errors in parentheses. The rightmost column provides a test that the overall effect of the given Blueprint element was zero in the given year. Rows that do not indicate a year refer to a test that the given Blueprint element had a zero effect in the base year, which was 2000-2001 for all Blueprint elements except for peer coaches, for which the base year was 1999-2000.

* Significant at 5%.

** Significant at 1%.

TABLE 18 Pooled models with year variations for high school students

	Model iv	Tests that overall effect equals 0 in a given year
Class size	Yes	
Teacher qualifications	Yes	
Interactions1:	No	
Block, Core, peer coach		
Interactions2:	No	
EDRP, summer		
Literacy Block	-0.0671	**
	(0.0153)**	
Literacy Block and 1999-2000 school year	0.0535	
	(0.0196)**	
Literacy Block and 2001-2002 school year	0.0268	**
	(0.0199)	
Literacy Block and 2002-2003 school year	-0.0039	**
	(0.0194)	
Literacy Block and 2003-2004 school year	0.0276	*
	(0.0193)	
Literacy Block and 2004-2005 school year	0.0914	
	(0.0305)**	
Literacy Core	-0.1380	**
	(0.0305)**	
Literacy Core and 2001-2002 school year	0.0283	**
	(0.0439)	
Literacy Core and 2002-2003 school year	0.0699	
	(0.0433)	
Literacy Core and 2003-2004 school year	-0.2463	*
	(0.1715)	
Literacy Block/Core for EL students	-0.2275	**
	(0.0182)**	
Literacy Block/Core for EL students and 1999-2000 school year	0.0371	**
	(0.0268)	
Literacy Block/Core for EL students and 2001-2002 school year	0.0403	**
	(0.0220)	
Literacy Block/Core for EL students and 2002-2003 school year	0.1065	**
	(0.0241)**	
Literacy Block/Core for EL students and 2003-2004 school year	0.1231	**
	(0.0259)**	

Continued on next page

Continued

Literacy Block/Core for EL students and 2004-2005 school year	0.2674	
	(0.0384)**	
Summer session	0.0053	
	(0.0209)	
Summer and 2001-2002 school year	-0.0276	
	(0.0249)	
Summer and 2002-2003 school year	-0.0296	
	(0.0248)	
Summer and 2003-2004 school year	-0.0261	
	(0.0258)	
Summer and 2004-2005 school year	0.0261	
	(0.0904)	
Average years experience of peer coach	-0.0011	
	(0.0003)**	
Peer coach as % of enrollment	0.0382	
	(0.0751)	
Peer coach and 1999-2000 school year	-0.2050	
	(0.1123)	
Peer coach and 2001-2002 school year	-0.1475	
	(0.2254)	
Peer coach and 2002-2003 school year	0.5059	**
	(0.1452)**	
Peer coach and 2003-2004 school year	-0.0219	
	(0.1694)	
Peer coach and 2004-2005 school year	1.1006	**
	(0.3611)**	
Observations	97656	
Number of stud id	53416	
R-squared	0.61	
P-values for F-tests that selected year interactions equal zero:		
All Blueprint variables	<0.0001	
Literacy Block	0.0042	
Literacy Core	0.1661	
Literacy Block/Core for EL	<0.0001	
Summer	0.7511	
Peer coach as % of enrollment	<0.0001	

NOTES: Standard errors in parentheses. The rightmost column provides a test that the overall effect of the given Blueprint element was zero in the given year. Rows that do not indicate a year refer to a test that the given Blueprint element had a zero effect in the base year, which was 2000-2001 for all Blueprint elements except for peer coaches, for which the base year was 1999-2000.

* Significant at 5%.

** Significant at 1%.

TABLE 19 Separate models by subject area of the number of A–G classes taken in a given year

	Foreign Language	Art	Math	English	Science	Social Studies
Total Block interventions	-0.65	-0.26	-0.09	0.27	-0.23	0.07
	(0.02)**	(0.01)**	(0.02)**	(0.01)**	(0.02)**	(0.02)**
Total Core interventions	-0.99	-0.41	0.19	0.47	-0.07	-0.20
	(0.03)**	(0.03)**	(0.03)**	(0.03)**	(0.03)	(0.04)**
Total Block/Core interventions for EL students	-0.72	-0.24	-0.04	-0.01	-0.18	0.12
	(0.02)**	(0.02)**	(0.02)*	(0.02)	(0.02)**	(0.02)**
Total summer interventions	-0.18	0.04	0.05	0.03	-0.07	0.07
	(0.02)**	(0.02)*	(0.02)**	(0.01)*	(0.02)**	(0.02)**
R^2	0.20	0.03	0.04	0.05	0.10	0.21
Observations	140,037	140,037	140,037	140,037	140,037	140,037

** Denotes significance at the 1 percent level of confidence.

* Denotes significance at the 5 percent level of confidence.

References

American Institutes for Research (AIR). 2002. *Evaluation of the Blueprint for Student Success in a Standards-Based System.* Palo Alto, CA: American Institutes for Research.

Ashenfelter, Orley. 1978. "Estimating the Effect of Training Programs on Earnings." *Review of Economics and Statistics* 60 (1): 47–57.

Bassok, Daphna, and Margaret E. Raymond. 2005. "Performance Trends and the Blueprint for Student Success." In *Urban School Reform: Lessons from San Diego*, ed. Frederick M. Hess (Cambridge, MA: Harvard Education Press), 299–323.

Betts, Julian R. 2005. "The Promise and Challenge of Accountability in Public Schooling." In *Urban School Reform: Lessons from San Diego*, ed. Frederick M. Hess (Cambridge, MA: Harvard Education Press), 157–176.

Betts, Julian R., Andrew Zau, and Kevin King. 2005. *From Blueprint to Reality: San Diego's Education Reforms.* San Francisco: Public Policy Institute of California.

Betts, Julian R., Andrew Zau, and Lorien Rice. 2003. *Determinants of Student Achievement: New Evidence from San Diego.* San Francisco: Public Policy Institute of California.

Coleman, James Samuel. 1966. *Equality of Educational Opportunity.* Washington, DC: U.S. Dept. of Health, Education, and Welfare, Office of Education.

Hightower, Amy M. 2002. "San Diego's Big Boom: Systemic Instructional Change in the Central Office and Schools." In *School Districts and Instructional Renewal*, ed. Amy M. Hightower, Michael S. Knapp, Julie A. Marsh, and Milbrey W. McLaughlin (New York: Teachers College Press), 76–93.

Quick, Heather E., Beatrice F. Birman, Lawrence P. Gallagher, Jean Wolman, Kassandra Chaney, and Hiroyuki Hikawa. 2003. *Evaluation of the Blueprint for Student Success in a Standards-Based System: Year 2 Interim Report.* Palo Alto, CA: American Institutes for Research.

Stein, Mary Kay, Lea Hubbard, and Hugh Mehan. 2004. "Reform Ideas That Travel Far Afield: The Two Cultures of Reform in New York City's District #2 and San Diego." *Journal of Educational Change* 5 (2): 161–97.

PPIC PUBLIC POLICY INSTITUTE OF CALIFORNIA

PUBLIC POLICY INSTITUTE OF CALIFORNIA
500 Washington Street, Suite 600
San Francisco, California 94111
phone: 415.291.4400
fax: 415.291.4401
www.ppic.org

PPIC SACRAMENTO CENTER
Senator Office Building
1121 L Street, Suite 801
Sacramento, California 95814
phone: 916.440.1120
fax: 916.440.1121