



Lessons in Reading Reform

Finding What Works

Julian R. Betts • Andrew C. Zau • Cory Koedel

Supported with funding from the Donald Bren Foundation



iStockphoto

SUMMARY

The San Diego Unified School District, the nation's eighth-largest, launched an ambitious program of literacy reforms in 2000 aimed at narrowing reading achievement gaps. Known as the Blueprint for Student Success, the program ran through 2005. The reforms succeeded in boosting the reading achievement of students who had been identified as lagging behind at the elementary and middle school levels. The key element that seems to have driven this success was a significant amount of extra student time spent on reading, with a possible collateral factor being widespread professional development for district teachers. The combination was neither cheap to implement nor a magic bullet. But in elementary and middle schools it demonstrably worked. In high schools, with one exception, it did not.

This study summarizes our statistical evaluation of all of the Blueprint reforms over the five-year period, drawing lessons for educators about why some elements of the Blueprint succeeded and how they could be implemented elsewhere. Elements that appeared particularly helpful were extended-length English classes in middle school and an extended school year for low-performing elementary schools. Even in high schools, we found that students who participated in triple-length English classes were more likely to be promoted to the next grade. There were several goals that the Blueprint interventions did not achieve. But neither did the interventions confirm the fears of many Blueprint detractors—such as that extra time spent on reading would degrade student performance in other subjects or

would cause student burnout, all to the detriment of their entire school careers. The Blueprint appeared to have little or no bearing on student success in completing high school college preparatory work.

One of the lessons of the Blueprint is that specific changes in both state and federal government policy could foster these kinds of ambitious reforms elsewhere, at the school district level. California could continue its recent trend of collapsing categorical funding into more flexible mechanisms that give individual school districts freedom for reforms that boost achievement in the most appropriate way. At the federal level, the Department of Education could ease its Title I waiver requirements, so that districts could use that money for reforms that target not only low-income students, but also low-performing students, regardless of school or neighborhood.

A key aspect of San Diego's reform program was that it was comprehensive and coherent. Interventions often were applied in two or more of the elementary, middle, and high school grade spans. Further, professional development was delivered uniformly, with a single focused goal, to teachers throughout the district. But perhaps the most important lesson for education policymakers is that many of the reforms took several years to bear fruit. Most notably, the peer coaching system for teachers did not typically generate positive gains in the first year or two, but did appear to do so by the later years. An obvious lesson here is that school district leaders everywhere, when they implement reforms, must show considerable patience in their quest for improved student literacy.

Please visit the report's publication page to find related resources:
<http://www.ppic.org/main/publication.asp?i=922>

Introduction

Beginning in the 1990s, a national movement to enhance the accountability of public school systems gathered strength and culminated in 2001 with the passage of federal No Child Left Behind (NCLB) legislation. NCLB formalized reforms that many states had already initiated, such as creating academic content standards, implementing statewide student testing systems, and developing systems of rewards and sanctions for schools based on student performance on state tests. California had implemented a similar accountability system of its own in 1999.

In early 2010, President Obama announced plans for the reauthorization of NCLB. The proposed legislation calls for better measures of student learning but maintains the original concepts of measuring student performance and intervening when necessary. Despite changes from the original, the new NCLB will continue to emphasize content standards, testing, rewards, and sanctions.

However, federal and state reforms have not yet presented many ideas about exactly *how* individual schools and school districts should intervene to help students who are lagging behind. NCLB calls for failing schools to implement tutoring and busing—and for schools that repeatedly fail to make adequate yearly progress to face a series of sanctions, up to and including removal of administrators. But it is not particularly prescriptive about how teachers should teach, how schools should organize the school day, or what curricula they should use to correct deficiencies. The accountability movement presupposes that we know how to help students who are struggling academically. In fact, the literature on the effects of specific reforms, such as professional development, reduced class size, and summer school, is quite mixed. Nor does the policy community have much evidence on how best to intervene when students and schools fail to meet standards.

These knowledge deficits increase the importance of careful evaluation of interventions that have already been attempted. Increasingly, individual school districts have become laboratories for interventions aimed at improving



AP PHOTO/MARCIO JOSE SANCHEZ

Policymakers have offered little guidance about how best to intervene when students and schools fail to meet standards.

the achievement of students who fare poorly on state-mandated tests. Chicago public schools, for example, implemented interventions for students with low achievement that received much attention in the policy world (and, to a slightly lesser extent, in academic circles).¹

Individual school districts have become laboratories for interventions aimed at improving the achievement of students who fare poorly on state-mandated tests.

The San Diego Unified School District (SDUSD) has also received national attention for a series of literacy reforms it implemented from 2000 to 2005, called the Blueprint for Student Success. This report presents a quantitative evaluation of the effects the Blueprint had on student achievement in San Diego. Our findings hold lessons for districts elsewhere.

Data and methods

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

Our main interests are gains in student scores on state-administered reading tests, of which there were two during the Blueprint reform period: the Stanford 9 test in spring 1998 through spring 2002, and the California Standards Test (CST) in spring 2002 and later years. Our models of test scores avoid comparisons among different students. They instead compare individual students' achievement growth in years and grades when they participated in Blueprint program elements with growth during years and grades when they did not.

We present all estimated effects in terms of the number of percentile points by which a student is estimated to have moved as a result of participating in a given intervention. This means that a student who improves from the 25th to the 27th percentile would have initially scored better than 25 out of every 100 students in his or her grade, and better than 27 out of every 100 students after an intervention. The greater the gap between the two percentiles, the greater the improvement.

To these models we added numerous characterizations of Blueprint elements. For example, we tested whether the intensity of peer coach support in a school—measured as the ratio of peer coaches to overall enrollment—influenced reading. Because class size varies little across schools in the district (Betts, Zau, and Rice 2003), a peer coach who had to work with a greater number of classrooms could be less effective. Because a peer 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 each school.

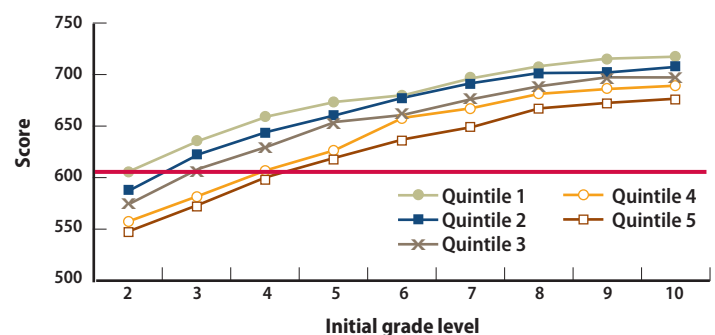
More detail on our data and methods appears in the online technical appendix, available at http://www.ppic.org/content/pubs/other/810JBR_appendix.pdf.

Reforms included extra time on task for students, including after-school and before-school interventions, along with summer programs.

Figure 1 illustrates just how large the achievement gaps were in San Diego before the Blueprint was implemented. The figure shows mean scores on the spring 1998 Stanford 9 reading test, dividing the district's schools into five categories (quintiles) according to socioeconomic status (measured according to the share of students eligible for federal meal assistance). It shows that students attending schools in the poorest quintile read at levels two to five grades behind those of students attending schools in the most affluent quintile. For example, the horizontal line shows that students attending schools in the most affluent quintile had test scores at the end of grade 2 that were not matched by students attending the poorest quintile of schools until they were partway through grade 5.

Facing such large achievement gaps, and in light of the large English Learner (EL) population in San Diego, in the spring of 2000 the district began implementing sweeping

Figure 1. In 1998, students in the lowest socioeconomic quintile of schools read at levels two to five grades behind those of students in the highest quintile



SOURCE: Betts, Zau, and Rice 2003.

NOTES: Students in the least affluent quintile were reading below the 2nd-grade reading level of students in the most affluent quintile (indicated by the red line) when tested in the spring of grade 4 and tested slightly above the line in the spring of grade 5. Similarly, the reading achievement levels reached by students in grades 3 to 5 in the most affluent quintile of schools is not reached until two to five grades later in the poorest quintile of schools.

reforms to boost English literacy. SDUSD superintendent Alan Bersin enlisted the help of chancellor of instruction Anthony Alvarado to develop and implement the Blueprint; Alvarado had experience and success with similar reforms as superintendent of Community School District #2 in New York City. Using federal Title I money and other funds derived in part from foundation grants, the district developed a series of interventions for students lagging behind in reading. Reforms included extra time on task for students, including after-school and before-school interventions, along with summer programs. The district also focused on professional development for teachers, in large part through the assignment of specially trained teachers—peer coaches—to every school in the district.

Resistance to the Blueprint reforms was quite strong from the beginning, and seemed to grow. The level of San Diego school board support for the reforms fell, and in mid-2005, the board voted not to renew Superintendent Bersin’s contract. Statewide financial cutbacks to education in the later years of the Blueprint also made it impossible for the district to continue the programs indefinitely. Other than double-length class known as Literacy Block, all Blueprint interventions were ultimately eliminated.²

When the concerns of Blueprint opponents are stated specifically, some can be tested empirically. For example, some worried that the unprecedented focus on reading and writing would draw attention away from student learning in other key areas, such as mathematics. Another concern,

When the concerns of Blueprint
opponents are stated specifically, some can
be tested empirically.

voiced strongly by some members of the Latino community, was that assigning high school students to double- or even triple-length English classes could hamper their ability to complete the A–G sequence of high school courses required for admission to the two California public univer-

sity systems (Ochoa 2001a, 2001b). Some critics went so far as to argue that the reforms would discourage students and prompt more of them to drop out of school entirely.

This report addresses these criticisms and presents the first analysis of whether the literacy reforms affected the completion of college preparatory courses or high school itself.³

Blueprint Reform Elements

The Blueprint that emerged in San Diego stressed the concept of balanced literacy, which emphasizes participation by students in reading, speaking, and writing, with the teacher initially actively supporting the students and then gradually demanding more of them as they progress.⁴ One of two overarching strategies was prevention—helping students whose reading skills were at or above grade level to reduce their chances of falling behind later. Some preventive measures were generally targeted, including extensive training of teachers and additional classroom materials. The second strategy was intervention: teachers identified students reading below grade level, and those students received extra instruction through four program elements.

There were four key prevention elements:

1. **Genre Studies.** A two-period English class for students in their first grade of middle or junior high school (either grade 6 or 7) and related professional development. The district considered this preventive because it reinforced the already strong skills of students who were reading at or above grade level.
2. **Peer coaching.** Each school was assigned at least one peer coach, to improve teaching.
3. **Focus schools.** The elementary schools with the weakest scores in the state test (the lowest tenth, or decile, statewide), received an extended school year, a second peer coach, and additional funds and staff.
4. **API-2 schools.** The elementary schools that ranked in the second-lowest decile in the state in the state Academic Performance Index (API) received a second peer coach and additional funds but did not extend the school year.⁵

Table 1. 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, during summer or intersession |

NOTE: English Learners were eligible for all of the interventions.

The intervention strategies of the Blueprint reforms were targeted at students whose test results indicated that they were reading below grade level. (These tests were different from the state-mandated tests discussed above.) In addition, EL students were strongly encouraged to participate in all of the interventions, regardless of their test scores. Schools placed EL students directly into extended-length English classes, but parents made final decisions on activities outside the regular school day, such as after-school or summer reading sessions. There were four key intervention elements:

1. **Literacy Block.** A double-length English class, this variant of Genre Studies was offered to students in grades 6–10 who lagged below or significantly below grade level, which the district determined using its own reading tests.
2. **Literacy Core.** For students significantly below grade level in grade 9, the English class was extended to three periods. In 2001–2002, grade 6 and 7 students also began to participate.
3. **Extended Day Reading Program (EDRP).** Students in grades 1 through 9 who were below and significantly below grade level participated in three 90-minute

periods of supervised reading each week, before or after school.

4. **Summer school and intersession.** Blueprint summer school was aimed at students in most grades from kindergarten through grade 9 who lagged in reading. Students were asked to attend for six weeks, four hours per day.⁶ Some schools, mostly elementary, have year-round schedules that did not permit the implementation of Blueprint summer school, so students in affected grades at these schools who lagged in reading participated in special intersession studies.

An additional program, grade retention (also known as accelerated classes), called for students significantly below grade level in their first year of elementary, middle, junior high, or high school (grades 1, 6, 7, or 9, respectively) to be held back a grade and then placed in classes that provided intensive (accelerated) remediation in literacy the next year. These were essentially Literacy Core classes. With only a few exceptions, this part of the Blueprint was not implemented, in part because California law gives individual teachers the final say on grade retention. However, a small number

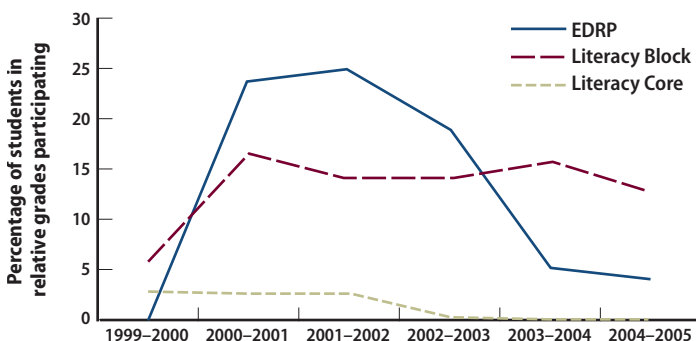
Most elements of the reform effort were implemented in 2000–2001 and expanded in 2001–2002.

of middle school students were put into this program, so we controlled for grade retention/accelerated classes in our middle school analyses.

In 1999–2000, the first year for which there was a test-score gain, the Blueprint was not in place, except for the peer coaching and Genre Studies elements, which were implemented on an extremely limited basis, and Literacy Block, which was implemented in grades 9 and 10. Most elements of the reform effort were implemented in 2000–2001 and expanded in 2001–2002. (The preventive Genre Studies classes were introduced first in grade 6 in 1999–2000, then in grade 7 in 2000–2001.) The reforms were scaled back between 2002–2003 and 2004–2005, and Literacy Core was canceled at the start of the 2003–2004 school year.

Figure 2 shows an average across relevant grades of the percentage of students participating in three different interventions—the Extended Day Reading Program (EDRP), Literacy Block, and Literacy Core. In all cases, there is a steep increase in participation followed by a decline, sometimes gradual and in other cases quite steep.

Figure 2. Statewide funding cuts contributed to the decline in Blueprint participation



NOTE: Average taken across relevant grades for each intervention. EDRP: grades 1 through 8; Literacy Block: grades 6 through 9; Literacy Core: grades 6, 7, and 9.



iStockphoto

At the elementary level, Focus and API-2 school prevention measures raised reading levels.

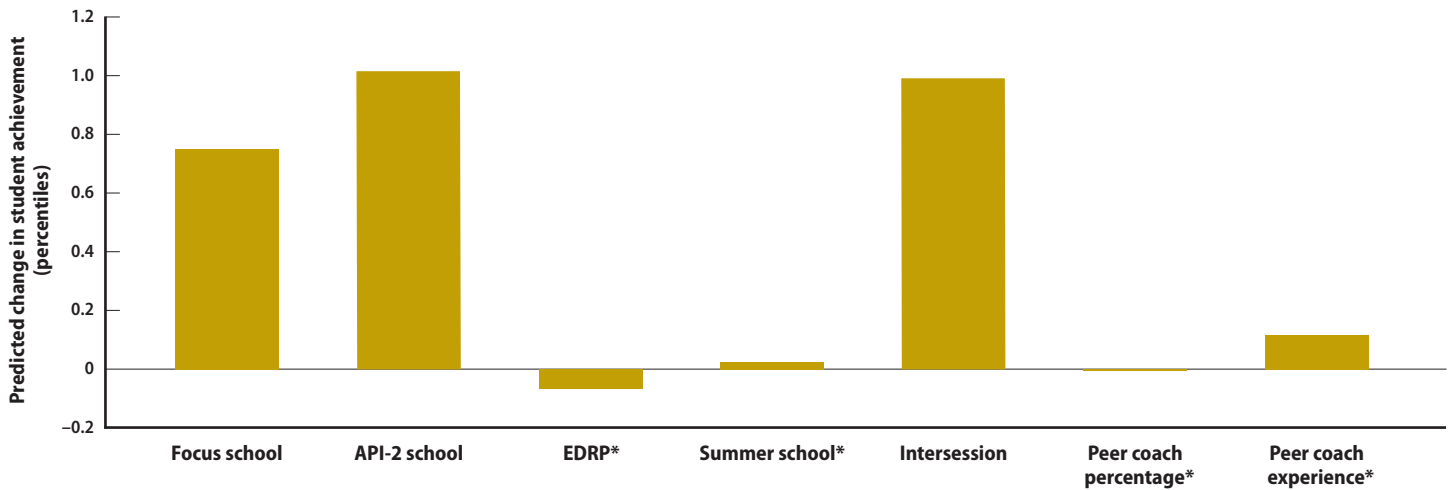
The Literacy Block element is something of an exception because it persisted for roughly four years after the formal end of the Blueprint in 2005.

Effects on Reading Test Scores

Figures 3, 4, and 5 show the main results for reading at the elementary, middle, and high school levels, respectively. In these figures, we show the estimated effects of participating in a given Blueprint intervention or preventive program. The two bars related to peer coaching have slightly different meanings: the first shows the estimated effect of increasing the number of peer coaches in a school as a percentage of enrollment by 0.1 percent, and the second shows the estimated effect of increasing by one the average years of teaching experience.⁷

At the elementary level (Figure 3), we found that a number of Blueprint interventions mattered, and that others had insignificant effects. Two preventive measures, the Focus and API-2 school reforms, both boosted reading gains. Students who attended these schools in the relevant years increased their reading performance by 0.75 and

Figure 3. Three Blueprint elements had positive effects on student achievement in elementary schools



NOTES: "Peer coach percentage" is the number of peer coaches expressed as a percentage of the number of students enrolled at the school; the bar shows the estimated effect of increasing this value by 0.1 percent. The bar for "Peer coach experience" shows the estimated effect of increasing by one the average years of teaching experience.

*Effects were not statistically significant at the 5 percent level and therefore should not be considered as significantly different from zero.

1.0 percentile points per year. These effects are considerable: a student who attended a Focus school for four years would be expected to move up in the district rankings by 3 percentile points. Another program element that made a difference was the intersession literacy program for students at year-round schools at which Blueprint summer school could not be held. These students moved up in the district rankings by about 1 percentile point per year more than they would have without the program.

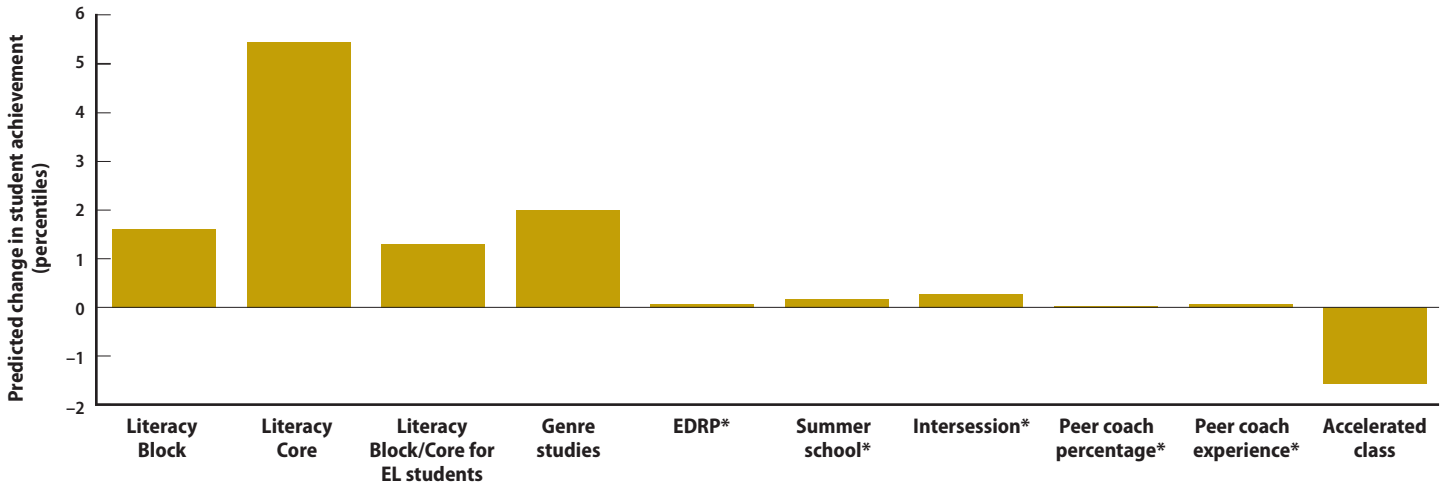
In contrast, EDRP and Blueprint summer school did not significantly affect students; nor did the ratio of peer coaches to students or the average level of experience of peer coaches.

In middle schools, what clearly stands out is that the extended-length English classes uniformly improved reading achievement (Figure 4). We estimate that the double-length (Literacy Block) and triple-length (Literacy Core) classes increased participants' reading rankings in the district by 1.6 and 5.5 percentile points, respectively, per year. These are very big shifts. At the end of three years, a student who enrolled in Literacy Core in grades 6 and 7, and then in Literacy Block in grade 8 (Literacy Core was offered in grades 6, 7, and 9 only), would be 12.6 percentile

points above the level he or she would have reached without the Blueprint interventions.

The value of more classroom time spent on reading is also evident for students who participated in the preventive double-length classes known as Genre Studies. Participants with reading skills at or above grade level saw their reading score rankings rise about 2 percentile points over the course of grade 7. This is roughly the same effect that we found for Literacy Block, the corresponding element for students reading below grade level. Finally, for EL students, the average effect of the double- and triple-length classes was also positive—their scores rose about 1.3 percentile points per year. However, students in accelerated classes (grade retention accompanied by triple-length English classes) experienced a 1.6 percentile point drop in their relative standing. But because the students held back a grade were also in Literacy Core, it is the sum of the Literacy Core (+5.5 percentile points) and accelerated class (–1.6 percentile points) effects that best summarizes their experience. Thus, they are predicted to have gained, overall, 3.9 percentile points during the year in which they were retained. The remaining Blueprint variables do not enter significantly.

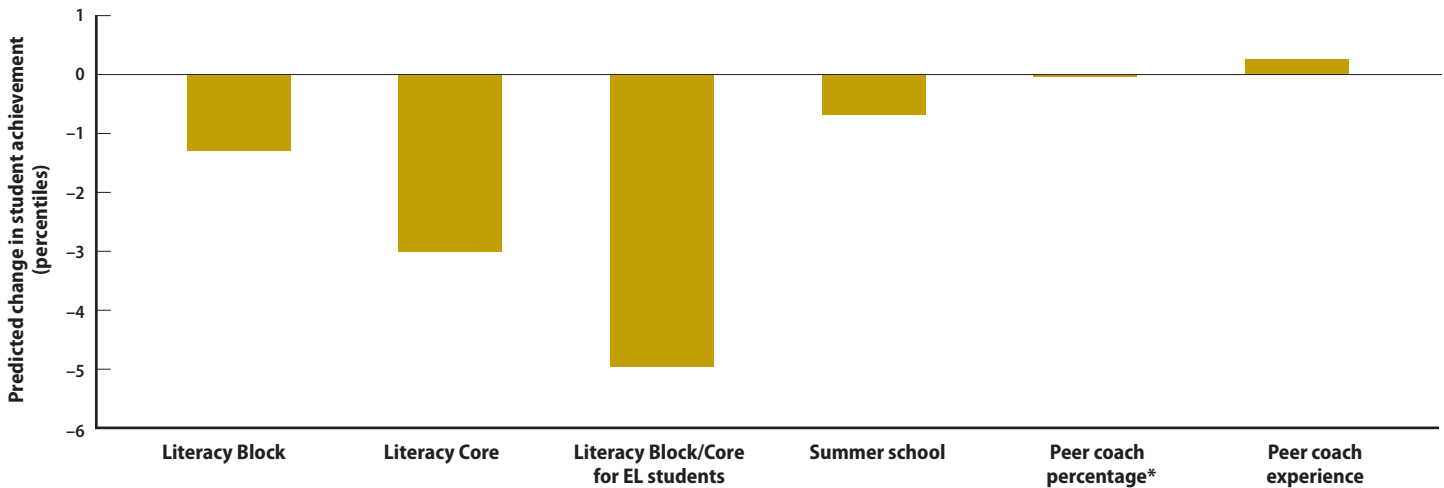
Figure 4. Blueprint strategies had a large effect on middle school student achievement



NOTES: "Peer coach percentage" is the number of peer coaches expressed as a percentage of the number of students enrolled at the school; the bar shows the estimated effect of increasing this value by 0.1 percent. The bar for "Peer coach experience" shows the estimated effect of increasing by one the average years of teaching experience.

*Effects were not statistically significant at the 5 percent level and therefore should be considered as insignificantly different from zero.

Figure 5. Most Blueprint elements had negative effects on student achievement in high schools



NOTES: "Peer coach percentage" is the number of peer coaches expressed as a percentage of the number of students enrolled at the school; the bar shows the estimated effect of increasing this value by 0.1 percent. The bar for "Peer coach experience" shows the estimated effect of increasing by one the average years of teaching experience.

*Effects were not statistically significant at the 5 percent level and therefore should be considered as insignificantly different from zero.

However, in high school, four Blueprint program elements appear to have negatively influenced reading achievement: Literacy Block, Literacy Core, summer session, and additional average years of teaching experience for peer coaches (Figure 5). Some of these effects are quite large.

Literacy Block/Literacy Core for high school EL students, for example, is associated with a drop of 4.9 percentile points for each year of student participation. For non-EL students, participation in double- and triple-length classes is associated with drops of 3.0 and 1.3 percentile points per



AP PHOTO/LISA HUDSON

Some Blueprint interventions may have a greater impact if teachers have more experience.

year, respectively. We discuss possible reasons for these negative results in the conclusion.

The Role of Teacher Experience

A number of studies have found evidence that teachers in their first few years of classroom experience are not as effective as more experienced teachers. The impact of Blueprint interventions might vary with teacher experience, but the direction of the effect is unclear. 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, the interventions could be more effective when teachers are more experienced.

We ran models that interacted teachers' experience levels with the various Blueprint indicators. In elementary schools, we focused on the homeroom teacher, while in middle schools and high schools, we focused on the English teacher. In elementary schools, results suggested that

students with more experienced teachers gain more from the peer coaching and API-2 Blueprint elements. (Increases in the peer-coach-to-enrollment ratio become statistically significant for more experienced teachers, but the effects are small.) We also found some evidence that students gained less from the EDRP and summer school interventions if they

In elementary schools, results suggested that students with more experienced teachers gain more from the peer coaching and API-2 Blueprint elements.

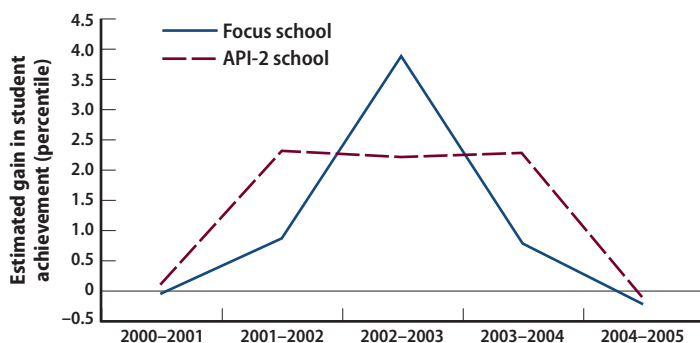
had teachers with relatively less experience—specifically, teachers with zero to two years of experience. The differences in effects are, in all cases, extremely small. At the middle school level, we found some evidence that Genre Studies and EDRP were less effective when the English teacher was relatively inexperienced, but these effects are also very small. Finally, at the high school level, we found no evidence that the English teacher's experience influenced the effect of the various Blueprint program elements.

Overall, the effects of the Blueprint typically did not vary with respect to teacher experience, but where we did find statistically significant effects, they suggest that the Blueprint elements were sometimes less effective when the students had less experienced teachers.

Program Effectiveness over Time

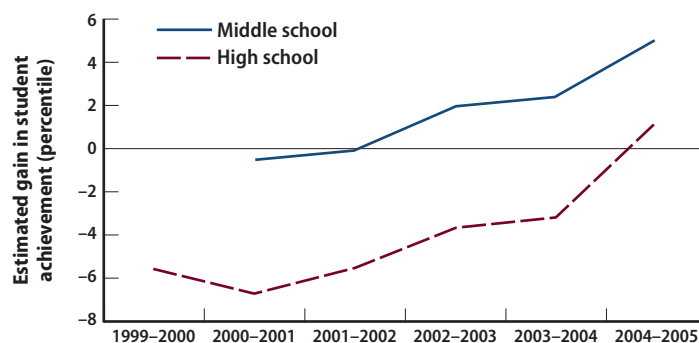
The most common pattern of Blueprint program effectiveness is one of rise and decline—it increased for one or more years and then began to fade. The Focus and API-2 programs in elementary schools, Genre Studies in middle schools, and Literacy Core for non-EL high school students all followed this pattern.⁸ The Focus effects exhibit a particularly steep rise and decline (Figure 6). Focus schools

Figure 6. Focus school and API-2 elements in elementary schools had initial positive effects that weakened over time



NOTE: For both types of schools, effects are significantly different from zero in 2001-2002 and 2002-2003.

Figure 7. Over time, Literacy Block and Literacy Core became more effective for EL students in middle school, and less ineffective in high school



NOTE: For middle schools, effects are significantly different from zero in 2002-2003 through 2004-2005; for high schools, effects are significantly different from zero for all years except 2004-2005.

had a longer school year (about 24 days) only in 2000-2001 and 2001-2002, after which financial constraints forced the cancellation of additional days. 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.⁹

We also found some evidence of gradual increases in effectiveness over time. Strikingly, in middle schools the effect of the Literacy Block/Core elements on EL students increased each year, without exception. In high schools, the effects of Literacy Block/Core for EL students also increased over time—that is, they became less negative each year, before becoming positive (but not significant) in 2004-2005. This suggests that middle and high school teachers (and their EL students) made better use of these extended-length English classes with each successive year (Figure 7). EL students appeared to gain substantially more from the Literacy Block/Core courses in middle school than in high school. This discrepancy is consonant with the finding by Zau and Betts (2008) that, on average, EL students redesignated as Fluent English Proficient (FEP) in the lower grades ultimately tend to do quite well on the California High School Exit Examination (CAHSEE) once they reach high school. However, EL students who have yet to be

redesignated as FEP by high school or who arrived in the United States in the high school years face a rough road to mastery of both English and the high school curriculum.

We also noted positive effects of peer coaching over time, although the pattern is not completely uniform.¹⁰ This is understandable, because it should take some time for peer coaches to visit classroom teachers, model teaching methods, and encourage adoption and mastery of the methods.¹¹

Objections to the Blueprint

Early opponents of the Blueprint, including a Latino coalition, worried that by placing so much emphasis on English-language literacy, the program would distract students from other subjects. This was especially likely to happen, it was argued, to underperforming students attending double- and triple-length English class.¹² It was also suggested that time spent on Blueprint activities would divert students from the completion of classes necessary for admission to the state's two public university systems. (Blueprint supporters countered by positing reading as a gateway skill that allows students to learn from textbooks in *all* subject areas.) A second concern of Blueprint opponents was that students would simply burn out from the additional time spent in longer classes and before- and after-school reading programs.



DAVID BUTOW/CORBIS

Blueprint literacy programs probably did not interfere with student performance or engagement in other subject areas.

We chose several ways to test these concerns. First, we assessed whether increased attention to English literacy resulted in deterioration of math skills, which are tested annually. Second, we looked to see whether student absences—an indication of burnout—increased because of participation. Finally, to measure whether overall academic progress might have slowed, we studied whether Blueprint participants were more likely to be retained a grade in the years in which they participated in various interventions. (As we noted earlier, the Blueprint program called for

grade retention at certain grade levels, but this element was never meaningfully implemented except in middle school, and even there only on a very limited basis.) At the high school level, we examined whether the Blueprint could have increased dropout rates or interfered with students' ability to complete the necessary course requirements for admission to California's public university systems.

We estimate that a few of the Blueprint program elements influenced math achievement, but the effects were usually very small and were roughly balanced between positive and negative. The largest negative effects appear in high school and suggest that participation in Literacy Block and Literacy Core is associated with a drop in participants' math achievement percentiles of about 1.3 and 1.8 points, respectively. Otherwise, Blueprint interventions did not divert students' attention strongly from learning math.

We looked to see whether student absences—
an indication of burnout—
increased because of participation.

In both elementary and high schools, we found that three Blueprint elements decreased absences and increased school attendance, and one element was associated with increased absences. This is again a mixed result, but one that, overall, is more supportive of the notion that the literacy programs did not systematically cause burnout. In middle schools, however, five Blueprint elements were found to be associated with increased absences, while three pointed the other way. We estimate that participation in Literacy Block or Literacy Core increased absences by about 0.3 and 0.4 percentage points, for non-EL and EL students, respectively. (By comparison, in high school, we found that these two programs would likely decrease absences by about 0.4 percent.) Given that the average student was absent 4.4 percent of the time in elementary school, 5.4 percent of the time in middle school and 4.8 percent of the time in high school, a shift of 0.3 or 0.4 points in either direction is a fairly big effect in relative terms. But in real terms, a 0.4 percent drop in attendance translates to less than one day out of a 180-day school year.

Were Blueprint participants more likely to be retained, that is, held back a grade? We found few associations in the elementary and middle school models to support that hypothesis. A few Blueprint variables appeared to matter, in both positive and negative directions, but the effects were quite small. In high schools, the effects were also small, but they are also uniformly negative; this suggests that Blueprint participation had the opposite effect in high school—it lowered the probability of being retained.¹³ The Blueprint element that had the largest effect on high school retention rates, in both directions, was the Literacy Core program. Students who attended were estimated to lower their probability of being retained by 3.6 percent. This is quite a large reduction, comparatively—on average, only 4.8 percent of students are retained a grade in high school. By another measure, we found indications that participating in one additional Blueprint intervention in high school lowered the probability of grade retention by 0.9 percent. Our overall conclusion is that the link between Blueprint participation and grade retention is weak and, typically, small and negative.

To evaluate the extent to which Blueprint interventions influenced whether students graduated from high school and whether they completed the course requirements for admission to California’s two public university systems, our analytical methods changed, in part because we were measuring longer-term outcomes and because graduation and course completion are unique events, not differences over time.¹⁴ We would expect a negative but noncausal relationship between Blueprint participation and these outcomes. Students identified as needing Blueprint interventions were, by definition, already having academic difficulties. They would therefore already be less likely than other students to graduate from high school and less likely to complete the necessary college preparatory coursework. It is important not to attribute the results of these built-in characteristics of the sample group to the Blueprint itself. Moreover, this selection problem limits our ability to make causal inferences.¹⁵

Our overall conclusion is that the link between Blueprint participation and grade retention is weak and, typically, small and negative.

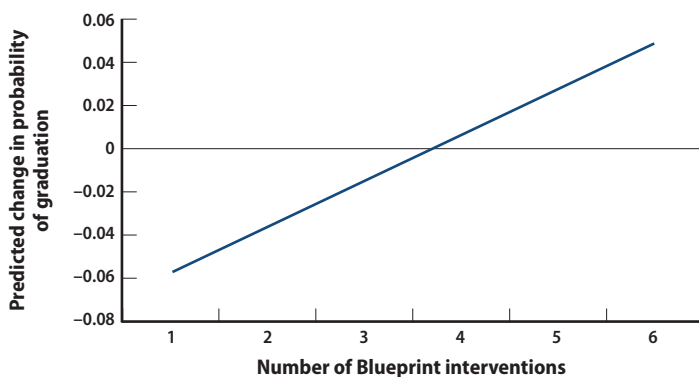
It is not surprising, therefore, that the raw data suggest that students who were involved in the Blueprint interventions during high school were much less likely to graduate. Students who had participated in at least one Blueprint intervention were about 9 percent less likely to graduate. When we control for characteristics of students and schools, this number drops to about 2 percent—still a large number in terms of the graduation rate, given that approximately 89 percent of the student sample ultimately graduated from high school. But again, we doubt that this means that the Blueprint program *caused* students to drop out; rather, their reading and writing deficiencies likely led directly to both their participation in the Blueprint and their lower probability of graduating.

We also tested the effects on graduation of participation in more than one Blueprint element.¹⁶ Figure 8 shows the results, which suggest a negative effect for those who took part in one or two interventions, roughly no effect for those who participated in three or four interventions, and a positive effect for those who took part in five or more interventions.

In a fashion similar to our graduation analysis, we evaluated the effects of Blueprint interventions on completion of courses necessary for admission to California's two state university systems, known as the A–G requirements. Again, we relied on the same set of student- and school-level variables to remove as much bias as possible from negative selection into the Blueprint program.

Approximately 36 percent of our entire student sample completed the university course requirements, and 29 percent of the total sample participated in at least one Blueprint intervention. When we examined completion rates in the raw data, we found that Blueprint participants were about 27 percent less likely than other students to complete all the A–G requirements.¹⁷ Again, this should not be interpreted as evidence of the Blueprint causing lower completion rates. Low achievers are more likely to be steered into the Blueprint supports and interventions in the first place. Just as with our graduation analysis, the relationship between Blueprint participation and A–G requirement

Figure 8. Students who participated in several Blueprint interventions were more likely to graduate from high school



NOTE: For middle schools, effects are significantly different from zero in 2002–2003 through 2004–2005; for high schools, effects are significantly different from zero for all years except 2004–2005.

completion becomes far smaller when we include student and school characteristics as variables. Participation in the Blueprint is associated with about an 8 percent reduction in the probability of completing the A–G requirements.

We also evaluated the effects of additional Blueprint interventions on the completion of the university course requirements only among students who participated in at least one Blueprint element, a technique that should lessen bias. We found that multiple Blueprint interventions apparently do not affect whether students complete the entire set of university course requirements, at least when we control for both student and school characteristics.

**Multiple Blueprint interventions
apparently do not affect whether students
complete the entire set of university
course requirements.**

We also examined how Blueprint interventions affect class-taking behavior on a year-by-year basis for each student, a more convincing approach. We inferred the effect of participation by comparing the number of A–G courses completed in years when the student participated in a Blueprint intervention to years when he or she did not.¹⁸

As one might expect, the summer school Blueprint 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 negative effects on student class-taking behavior within years, and are associated with a roughly one-for-one reduction in the number of university course requirements taken by students in the year of the intervention.

At first glance, these strong and negative results appear to be at odds with the previous finding that participation in additional Blueprint interventions does not affect whether a student completes *all* of the A–G requirements. Looking more closely, we can see that since Blueprint participants

were, on average, so far from completing the A–G course requirements, taking an extra class or two of English did little or nothing to lower the already low probability that they would complete all of them. Table 2 provides summary statistics for Blueprint participants’ progress in completing the university course requirements. It shows that almost three-quarters of the students participating in the Blueprint interventions ended up three or more subject requirements away from completing the full A–G set. Because each of the subject requirements requires a student to pass two to eight semester-length courses, we infer that most of these students would have fallen far short of completing the California public university admission requirements with or without the Blueprint reforms.

Finally, we examined the nature of the class-substitution behavior of Blueprint participants. Students who give up classes in order to participate in Blueprint most commonly drop foreign-language classes. They also tend to drop art and science classes. Blueprint participants are more likely to complete university-required courses in

Almost three-quarters of the students participating in the Blueprint interventions ended up three or more subject requirements away from completing the full A–G set.

English and, to some degree, in math and social studies. This implies that Blueprint interventions in fact encourage participants to take additional classes in the two A–G subjects that they are least likely to complete, English and math. Because these are core subjects, this could be construed as a salutary outcome, regardless of a participant’s post-high school plans.¹⁹

Conclusions

Our findings validate the idea that extra time on task for students who are behind in reading can lead to meaningful gains in literacy. The Blueprint reforms 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. It is clear that for elementary and middle school students, additional time on task—whether through double- or triple-length classes or longer school days—generally boosted reading achievement. The most impressive effect was from the Literacy Core program for non-EL students in middle school, which was associated with a rise of 5.5 percentile points in reading in the year the student participated—a sizable gain.

An important aspect of the Blueprint was professional development for teachers, operating in part through the placement of peer coaches in each school.²⁰ We found no effects when the ratio of peer coaches to enrollment at each school varied, 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 estimate that the overall effect of peer coaches was positive

Table 2. Most Blueprint participants were well short of completing A–G admission requirements

| | Share of Blueprint participants |
|---|---------------------------------|
| Number of required courses completed | |
| All | 13.7 |
| 5/7 | 12.7 |
| 4/7 | 12.4 |
| 3/7 | 14.6 |
| 2/7 | 16.2 |
| 1 or 0 | 30.5 |
| Subject areas completed | |
| Math | 32.4 |
| English | 35.0 |
| Science | 45.0 |
| Social studies | 72.0 |
| Art | 57.5 |
| Foreign language | 39.0 |

NOTE: The seventh university course requirement is 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 six listed in the table are completed.

if we take into account the possibility that it could vary with the experience of the classroom teacher. Specifically, we found that elementary students with more experienced teachers were more likely than students with less experienced teachers to gain from increases in the intensity of peer coaching at their schools.

The fact that extensive professional development for teachers accompanied extended-length classes should not be overlooked. We cannot test whether this professional development was crucial, because all teachers received this training. Without evidence on whether the teacher training was a prerequisite for the effectiveness of extended-length classes, a fiscally cautious policymaker might conclude that a district should invest in both extended-length classes and professional development.

High school adolescents might be particularly vulnerable to the stigma of being pulled out of regular English classes to participate in double- or triple-length English classes.

In sharp contrast to the results for lower grades, reforms at the high school level appear to have actually slowed gains in reading achievement for participants. It is impossible to know for sure why this was the case, but there are numerous possible explanations. The reforms in the lower grades built on Chancellor Alvarado's experience in implementing somewhat similar reforms primarily in kindergarten through grade 8 in New York's District #2 (Stein, Hubbard and Mehan 2004). This suggests that SDUSD administrators as a whole had little experience in implementing similar reforms at the high-school level. This would be an unfortunate explanation, since it would imply that reforms were dependent on the career paths and experiences of individual administrators—leaving in doubt the replicability of such reforms in districts without

similarly experienced administrators. Another possibility is that administrators had difficulty recruiting high school teachers for the crucial peer coaching positions. Third, as originally hypothesized by Betts, Zau, and King (2005), high school English teachers did not embrace the literacy reforms in the same way that teachers in lower grades did, in part because they viewed themselves as teachers of literature, not basic reading skills. Several teachers reaffirmed this attitude in conversations with us. Finally, Steele's theory of stereotype threat, in which students underperform when placed in situations in which they feel stereotyped, could play a role here (Steele 1997). Betts, Zau, and King (2005) hypothesize that high school adolescents might be particularly vulnerable to the stigma of being pulled out of regular English classes to participate in double- or triple-length English classes. One parent told the authors that, in her daughter's high school, students in the double- and triple-length English classes were resented by others for allegedly taking school resources away from other students; these students were often referred to in the hallways as "tards." One can imagine the effect of such epithets on a psychologically fragile adolescent.²¹

As for concerns about student burnout, we found no evidence that Blueprint participation increased students' rate of grade retention in any grade span. In fact, at the high school level, it appears that students who participated in many interventions were less likely to repeat a grade than in years when they did not participate. Similarly, Blueprint participation is related to *reduced* rates of absence at the high school level (although it is not systematically related to absences in lower grades). By these two measures, the Blueprint may have had positive effects at the high school level in spite of the negative effects on reading scores. As our final measure of burnout, we studied the decision to drop out of high school. Although it is true that, overall, Blueprint participants were more likely to drop out, those who participated extensively were significantly less likely to drop out than otherwise similar students who did not participate in the interventions at all.

This report has not focused on the costs of the Blueprint reforms. The expenditures, which the American Institutes

State policymakers may want to encourage districts to find ways to develop coherent and integrated interventions that span elementary, middle, and high schools.

for Research estimates at \$57.5 million in the 2000–2001 school year, were largely financed internally and through a waiver the district obtained to use federal Title I money to fund part of the reforms.²² Betts, Zau, and King (2005, 3–4) report that, over the five years of the Blueprint reforms, the district obtained \$33.5 million from three charitable foundations.²³ Clearly, these funds were only a small portion of the overall Blueprint expenditures over five years. Given an average of about \$6.5 million a year in external funding spread over about 130,000 students—an increase in spending of about \$50 per pupil per year—the capacity of some reforms to boost student reading achievements by several percentile points in lower grades is a point policymakers may wish to note.

Lessons and Recommendations

Four policy lessons emerge from our analysis.

1. Providing additional time for reading to students who are struggling, in a structured setting in which teachers have received training on teaching literacy, can indeed boost students' literacy levels. The effects of extended-length English classes for middle school students in San Diego were especially large.
2. Extending the school year at the lowest-decile elementary schools (Focus schools) may have been the main reason why this program boosted achievement so significantly. Replication of an extended school year in low-performing elementary schools, with and without the additional literacy supports provided by the Blueprint, would be a highly valuable exercise in other districts.
3. The Blueprint reforms countered the tendency, at least in California, to implement a slew of interventions in isolation from each other, creating an uncoordinated

and potentially incoherent overall approach to boosting achievement. California's laws have created hundreds of programs that provide state funding for specific K–12 programs. The Blueprint reforms challenged this approach, emphasizing an integrated and coherent strategy focused on literacy—professional development focused on providing teachers with similar skill sets across all grades, and a system of continuous literacy testing also encompassed all grades. Some Blueprint interventions were offered in all three grade spans.

4. Early intervention to aid students who lag behind in reading might be far more effective than intervention in high school. This finding in relation to the San Diego reforms could be of considerable importance to the 22 states that currently require students to pass a high school exit examination to obtain a high school diploma. In their study of the exit exam CAHSEE, Zau and Betts (2008) make a similar argument for intervening early in children's school careers, based on the finding that grade 4 test scores and report cards successfully predict which elementary school students will pass the high school exit exam six to eight years later. Interventions in grade 12 for students who had yet to pass California's exit exam seem to produce little gain.

State policymakers may want to encourage districts to find ways to develop coherent and integrated interventions that span elementary, middle, and high schools. One tactic would be to provide districts with more flexibility on how they spend state dollars. In fact, the state has moved in this direction in recent years, but explicit policies to encourage districts to implement reforms on a multi-grade basis would also be helpful.

Reforms at the federal level could also help districts implement comprehensive reforms. The federal govern-

ment could facilitate district innovation by making it simpler to apply for a waiver to use Title I money to subsidize district-wide interventions, on the condition that such interventions are carefully designed and aligned with the federal government's overall goal of boosting achievement across the board.²⁴

It is clear that many of the Obama administration's education reforms have to do with how school quality is measured. On the question of how states should intervene in schools that fail to make adequate yearly progress, the public has seen very few proposed changes from the various broad prescriptions outlined in the original NCLB. This

is not an accident. Rigorous evaluations have yet to reveal much about the best ways to help struggling students.

This study provides long-term evidence on one such intervention in San Diego. We have an acute need in the policy community to replicate and extend this and similar interventions in other locales. Combined with rigorous quantitative plans on how to study the impacts on student achievement of each intervention, such reforms could do much to help the accountability movement fulfill its original goal of not simply measuring student achievement but of acting decisively and effectively to remedy achievement gaps wherever they are found within a school district. ●

Acknowledgments

The authors thank many administrators at San Diego Unified School District, in particular Karen Bachofer (now at the University of California, San Diego) and Peter Bell, for helpful conversations and assistance with data issues. We are grateful to Debbie Beldock, Richard Murnane, David Neumark, Hans Johnson, and Lynette Ubois for their many useful suggestions. We thank Richard Greene and James Torr for skillful editorial assistance. We thank The William and Flora Hewlett Foundation, the Bill and Melinda Gates Foundation, and the Donald Bren Foundation for their generous financial support, without which this project would not have been possible. The data used in this project build upon a large database that a team led by Julian Betts has been working on since 2000. We warmly acknowledge funders of our previous projects that created the data infrastructure that made the current project feasible: The Atlantic Philanthropies, the Bill and Melinda Gates Foundation, the Girard Foundation, The William and Flora Hewlett Foundation, and the Public Policy Institute of California.

A technical appendix to this report is available on the PPIC website:
http://www.ppic.org/content/pubs/other/810JBR_appendix.pdf

Notes

¹ For examples of academic studies, see Bryk 2003 and Jacob 2003.

² One of the strongest sources of opposition was teachers, who criticized what they saw as top-down imposition of prescriptive reforms and the implication—in the establishment of peer coaching and other forms of professional development—that their knowledge and experience were being ignored. Ravitch reports that teachers “uniformly were bitter about the high-handed way in which the reforms were imposed on them. . . . Those who didn’t go along were bullied” (Ravitch 2010). Although these claims may be valid, the point is to know whether the reforms succeeded in boosting student achievement.

³ The present report builds directly on the work of Betts, Zau, and King (2005), which represents the first and only student-level analysis of the impact of the reforms. 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 in fact 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 probability of grade retention, the completion of the college-preparatory sequence in high school, and the completion of high school. See pages 7–10 of Betts, Zau, and King 2005 for an overview of related work.

⁴ For more details see Betts, Zau, and King 2005, and especially Stein, Hubbard, and Mehan 2004.

⁵ API is the acronym for the Academic Performance Index, a statistic measuring overall student achievement in a school. The California Department of Education calculates the API for each school annually. It also ranks schools into ten API deciles. Hence API-2 schools rank in the second-lowest decile of achievement statewide.

⁶ In addition, all secondary school students with D/F grades attended a more traditional type of summer school consisting of six weeks of courses in core subjects.

⁷ By comparison, the average number of peer coaches as a percentage of enrollment ratio was 0.2, 0.1, and 0.05 percent in elementary, middle, and high schools, respectively, and the average years of peer coaches’ teaching experience was 14, 12, and 14.

⁸ The last of these is slightly different in that the overall effect never becomes positive and significant.

⁹ We thank Karen Bachofer for this insight.

¹⁰ For instance, in middle schools, the peer coach variable had a negative and significant effect in 1999–2000 and no significant effect in 2000–2001, followed by positive and significant effects in 2001–2002 and 2002–2003, before becoming statistically insignificant (but still positive) in the following two years.

¹¹ The effects of Blueprint elements may have varied in other ways over time. 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 grades 9 and up. Data availability limits our ability to test these possibilities. 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 year increased achievement gains in the current year. This occurred in just over half the cases. In only about 10 percent of cases was there 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. We conclude that, in elementary and middle schools, the Blueprint led, in many cases, to both immediate and future gains in achievement; in high schools, the negative effects sometimes spilled over into the year following.

¹² Alberto Ochoa, co-chair of the San Diego County Latino Coalition on Education, in an October 29, 2001, op-ed in the *San Diego Union-Tribune*, expressed concern that Latinos would be predominantly assigned to the extra-length English classes. In a separate, October 9, 2001, letter on behalf of the coalition to the district school board, Ochoa equated the double- and triple-length English classes that the final version of the Blueprint implemented with academic tracking, which he argued would reduce Latinos’ ability to complete course requirements needed for admission to the University of California and California State University (UC and CSU) university systems. See Ochoa 2001a, 2001b.

¹³ In SDUSD, high school students are not formally retained. Rather, students’ grade level is determined by the number of semester credits earned (i.e., courses passed) to date. For the purpose of this report, the term “retained” indicates that a

student had not earned enough credits in a given school year to be considered “on-track” with his/her cohort/class.

¹⁴ In evaluating the extent to which Blueprint interventions affected whether students graduated from high school and whether they completed the course requirements for admission to California’s two public university systems, we cannot use our previous technique of student fixed effects to remove variation in ability or motivation across students. 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 bias as possible. At the student level, these variables include indicators for race, gender, EL status, parental education levels, and students’ standardized test scores in math and reading (Stanford 9) at the end of grade 8. The technical appendix (available at http://www.ppic.org/content/pubs/other/810JBR_appendix.pdf) describes the school-level variables we also included in these models.

¹⁵ 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 inferior approach, we compared one student with 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 a positive to a negative effect that is about 50 percent larger. Summer school for elementary students, instead of having zero effect on reading gains, 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 with another.

¹⁶ We can somewhat mitigate the negative 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. Nonetheless, these estimates could be biased downward or even upward. (A downward bias would arise if it were the truly struggling students who participated in more than one intervention. Less intuitively, an upward bias could arise if it were the more highly motivated students among those below grade level who elected to participate in more Blueprint interventions.)

¹⁷ There are seven A–G requirements. Students must take three years of college preparatory mathematics courses, four years of English, three years of history/social science, two years of laboratory science, two years of a foreign language, one year of visual and performing arts, and one year of additional college preparatory elective courses in any of these subjects.

¹⁸ Because the effects of Blueprint interventions on the completion of the university course requirements are likely to be most severe when they take away class time during the school year, in our model, we control separately for the four main interventions at the high school level: Literacy Block, Literacy Core, Literacy Block/Core for EL students, and Blueprint summer school.

¹⁹ Rose and Betts (2004) find evidence that passing certain high school math courses is highly indicative of success in college and the labor market a decade after graduation.

²⁰ The district also provided more traditional professional development to teachers on literacy topics. However, no consistent data on who participated in these other forms of professional development were kept, and so we focus on the peer coaches.

²¹ Steele (1997) argues that placing a student in a situation that emphasizes negative stereotypes about that student’s group can lead to underperformance.

²² American Institutes for Research 2002, page VII–7. This works out to roughly \$400 per student enrolled in the district that year. However, as we note later in the section, most of these funds do not represent additional district costs but rather a reallocation of expenditures.

²³ The foundations were the Bill and Melinda Gates Foundation, The Walter and Flora Hewlett Foundation, and the Atlantic Philanthropies.

²⁴ San Diego Unified had to apply to the federal government for a waiver to allow it to use federal Title I funding to help pay for the Blueprint reforms in Title I schools. Federal policy dictates that Title I money (for schools serving disadvantaged students) cannot supplant district spending for programs that already exist in all district schools. SDUSD had to apply for a waiver because the Blueprint’s theory of action was that any child, rich or poor, could lag behind in reading and writing and therefore could benefit from additional time on literacy tasks with the help of highly trained teachers.

References

- American Institutes for Research. 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.
- 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), 157–76.
- 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.
- Bryk, Anthony S. 2003. "No Child Left Behind, Chicago Style." In *No Child Left Behind? The Politics and Practice of School Accountability*, ed. Paul E. Peterson and Martin R. West (Washington, DC: Brookings Institution Press), 242–68.
- Jacob, Brian A. 2003. "A Closer Look at Achievement Gains Under High-Stakes Testing in Chicago." In *No Child Left Behind? The Politics and Practice of School Accountability*, ed. Paul E. Peterson and Martin R. West (Washington, DC: Brookings Institution Press), 269–91.
- Ochoa, Alberto. 2001a. Op-ed. *San Diego Union-Tribune*, October 26.
- Ochoa, Alberto. 2001b. Letter to SDUSD Board Members on behalf of the San Diego County Latino Coalition of Education, October 9.
- Ravitch, Diane. 2010. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic Books.
- Rose, Heather, and Julian R. Betts. 2004. "The Effect of High School Courses on Earnings." *Review of Economics and Statistics* 86 (2): 497–513.
- Steele, Claude M. 1997. "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist* 52 (6): 613–29.
- 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.
- Zau, Andrew C., and Julian R. Betts. 2008. *Predicting Success, Preventing Failure: An Investigation of the California High School Exit Exam*. San Francisco: Public Policy Institute of California.

About the Authors

Julian Betts is an adjunct fellow and a 2009–2010 Bren fellow at the Public Policy Institute of California. He is also professor and chair of economics at the University of California, San Diego, and a research associate at the National Bureau of Economic Research. He has written extensively on the link between student outcomes and measures of school spending, and he has studied the role that educational standards, accountability, teacher qualifications, and school choice play in student achievement. He has served on two National Academy of Sciences panels, the Consensus Panel of the National Charter School Research Project, and various advisory groups for the U.S. Department of Education. He is also principal investigator for the federally mandated National Evaluation of Magnet Schools. He holds a Ph.D. in economics from Queen's University, Kingston, Ontario, Canada.

Andrew C. Zau is a senior statistician in the Department of Economics at the University of California, San Diego. Previously, he was a research associate at the Public Policy Institute of California. Before joining PPIC, he was an SAS programmer and research assistant at the Naval Health Research Center in San Diego, where he investigated the health consequences of military service in Operations Desert Shield and Desert Storm. He holds a B.S. in bio-engineering from the University of California, San Diego, and an M.P.H. in epidemiology from San Diego State University.

Cory Koedel is an assistant professor of economics at the University of Missouri, Columbia. His research focuses on issues related to teacher quality, school choice, and curricular effectiveness. In 2005, he received the Spencer Foundation's prestigious Dissertation Fellowship Award and in 2008 was awarded the Outstanding Dissertation Award in Education Policy from the American Educational Research Association. He holds a Ph.D. in economics from the University of California, San Diego.

Board of Directors

WALTER B. HEWLETT, CHAIR
Director
Center for Computer Assisted Research
in the Humanities

MARK BALDASSARE
President and CEO
Public Policy Institute of California

RUBEN BARRALES
President and CEO
San Diego Chamber of Commerce

JOHN E. BRYSON
Retired Chairman and CEO
Edison International

GARY K. HART
Former State Senator and
Secretary of Education
State of California

ROBERT M. HERTZBERG
Partner
Mayer Brown LLP

DONNA LUCAS
Chief Executive Officer
Lucas Public Affairs

DAVID MAS MASUMOTO
Author and farmer

STEVEN A. MERKSAMER
Senior Partner
Nielsen, Merksamer, Parrinello, Mueller &
Naylor, LLP

CONSTANCE L. RICE
Co-Director
The Advancement Project

THOMAS C. SUTTON
Retired Chairman and CEO
Pacific Life Insurance Company

CAROL WHITESIDE
President Emeritus
Great Valley Center

PPIC is a private operating foundation. It does not take or support positions on any ballot measures or on any local, state, or federal legislation, nor does it endorse, support, or oppose any political parties or candidates for public office. PPIC was established in 1994 with an endowment from William R. Hewlett.

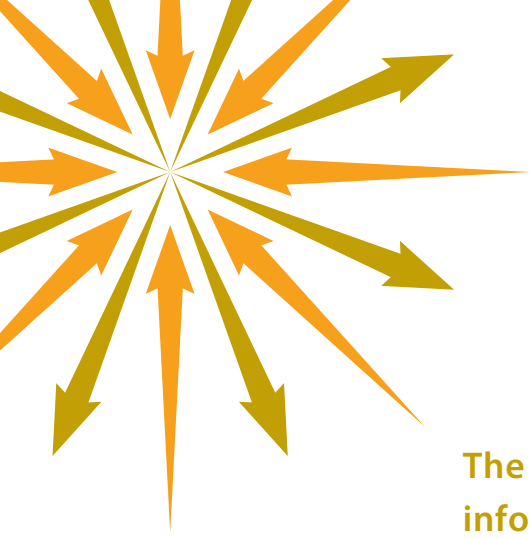
© 2010 Public Policy Institute of California. All rights reserved. San Francisco, CA

Short sections of text, not to exceed three paragraphs, may be quoted without written permission provided that full attribution is given to the source and the above copyright notice is included.

Research publications reflect the views of the authors and do not necessarily reflect the views of the staff, officers, or Board of Directors of the Public Policy Institute of California.

Library of Congress Cataloging-in-Publication Data are available for this publication.

ISBN 978-1-58213-140-5



The Public Policy Institute of California is dedicated to informing and improving public policy in California through independent, objective, nonpartisan research.

Additional resources related to education policy are available at www.ppic.org.



PPIC

PUBLIC POLICY
INSTITUTE OF CALIFORNIA

PUBLIC POLICY INSTITUTE OF CALIFORNIA
500 Washington Street, Suite 600 • San Francisco, California 94111
Telephone 415.291.4400 • Fax 415.291.4401

PPIC SACRAMENTO CENTER
Senator Office Building • 1121 L Street, Suite 801 • Sacramento, California 95814
Telephone 916.440.1120 • Fax 916.440.1121

