



## Green Dot Public Schools

### Intervention Description<sup>1</sup>

*Green Dot Public Schools* is a nonprofit organization that operates more than 20 public charter middle and high schools in California, Tennessee, and Washington. The *Green Dot Public Schools* are regulated and monitored by the local school district, but operate outside of the district's direct control. The *Green Dot Public Schools* model emphasizes high quality teaching, strong school leadership, a curriculum that prepares students for college, and partnerships with the community. Any student may enroll in a *Green Dot Public School* if there is space available. Many *Green Dot Public Schools* operate with unionized teachers and staff. Several of the *Green Dot Public Schools* were chartered in existing public schools which were performing below district or community expectations. Funding for *Green Dot Public Schools* operations comes through public federal, state, and local finances, while some transformations of existing district-run schools into charter schools have been funded partly by private foundations.

### Research<sup>2</sup>

The What Works Clearinghouse (WWC) identified one study of *Green Dot Public Schools* that both falls within the scope of the Charter Schools topic area and meets WWC group design standards with reservations. This study included 2,446 students in grades 9–12 in one school district in the United States. No studies meet WWC group design standards without reservations.

According to the WWC review, the extent of evidence for *Green Dot Public Schools* on the educational outcomes of students was small for four outcome domains—mathematics achievement, student progression, school attendance, and English language arts achievement. No studies meet WWC group design standards in the six other domains, so this intervention report does not report on the effectiveness of *Green Dot Public Schools* for those domains.<sup>3</sup> (See the Effectiveness Summary on p. 5 for more details of effectiveness by domain.)

### Effectiveness

*Green Dot Public Schools* had potentially positive effects on mathematics achievement, student progression, school attendance, and English language arts achievement for high school students.

### Report Contents

Overview	p. 1
Intervention Information	p. 3
Research Summary	p. 4
Effectiveness Summary	p. 5
References	p. 8
Research Details for Each Study	p. 9
Outcome Measures for Each Domain	p. 12
Findings Included in the Rating for Each Outcome Domain	p. 13
Supplemental Findings for Each Outcome Domain	p. 17
Endnotes	p. 23
Rating Criteria	p. 24
Glossary of Terms	p. 25

This intervention report presents findings from a systematic review of *Green Dot Public Schools* conducted using the WWC Procedures and Standards Handbook (version 3.0) and the Charter Schools review protocol (version 3.0).

**Table 1. Summary of findings<sup>4</sup>**

Outcome domain	Rating of effectiveness	Improvement index (percentile points)		Number of studies	Number of students <sup>5</sup>	Extent of evidence
		Average	Range			
<b>Mathematics achievement</b>	Potentially positive effects	+13	na	1	2,103	Small
<b>Student progression</b>	Potentially positive effects	+10	na	1	1,024	Small
<b>School attendance</b>	Potentially positive effects	+5	na	1	1,578	Small
<b>English language arts achievement</b>	Potentially positive effects	+4	na	1	1,865	Small

na = not applicable

### Intervention Information

#### Background

*Green Dot Public Schools* was founded in 1999 and continues to operate both start-up charter schools and turn-around charter schools. Address: 1149 S. Hill Street, Suite 600, Los Angeles, CA 90015. Web: greendot.org. Telephone: (323) 565-1600.

#### Intervention details

The *Green Dot Public Schools* model creates small community high schools and emphasizes a set of six principles. The first is to recruit, train, and support teachers who can accelerate student learning and close the achievement gap. The second is to offer autonomous school leadership to principals and assistant principals who are responsible for identifying high-quality teachers and providing them with personalized coaching and professional support. The third is to promote a college-going culture and offer a college preparatory curriculum. The fourth is to provide a range of comprehensive supports, including health and wellness services, gang intervention and mentorship programs, and college and career services. The fifth is to engage parents through adult education programming and self-advocacy training. Lastly, the *Green Dot Public Schools* model aims to be replicable in similar schools serving low-income students with unionized teacher workforces and similar funding sources.

Some schools in the *Green Dot Public Schools* network were chartered in existing public schools which were performing below district or community expectations. The study summarized in this report examined a group of *Green Dot Public Schools* formed by dividing an existing public high school into a set of smaller schools.

#### Cost

*Green Dot Public Schools* receive public funding to cover day-to-day operations. Transformation operations that transition traditional schools into the control of *Green Dot Public Schools* might have additional costs covered largely by private foundations. For example, *Green Dot Public Schools* spent approximately \$15 million over the course of a 4-year turnaround of a single high school campus in Los Angeles (Dillon, 2010), mainly from private sources.

## Research Summary

The WWC identified one eligible study that investigated the effects of *Green Dot Public Schools* on the academic outcomes of students. One additional study was identified but did not meet eligibility criteria (see the Glossary of Terms in this document for a definition of this term and other commonly used research terms) for review in this topic area. Citations for these two studies are in the References section, which begins on p. 8.<sup>6</sup>

The WWC reviewed one eligible study against group design standards. This study uses a quasi-experimental design that meets WWC group design standards with reservations. This report summarizes this one study.

**Table 2. Scope of reviewed research**

<b>Grade</b>	9–12
<b>Delivery method</b>	Whole school
<b>Intervention type</b>	Program

## Summary of studies meeting WWC group design standards without reservations

No studies of *Green Dot Public Schools* met WWC group design standards without reservations.

## Summary of studies meeting WWC group design standards with reservations

Herman et al. (2012) conducted a quasi-experimental study that examined the effects of *Green Dot Public Schools* on students in 11 high schools in the Los Angeles Unified School District (LAUSD), comparing students from eight *Green Dot Public Schools* to students from three comparison schools. This study meets WWC standards with reservations because it uses a quasi-experimental design in which the analytic intervention and comparison groups are shown to be equivalent. The study included three cohorts of students entering ninth grade and measured outcomes for up to 4 years. The authors compared students starting ninth grade in *Green Dot Public Schools* with similar students who started high school in the three comparison schools, matched using propensity scores based on gender, ethnicity, parents' education, poverty status, language classification, eighth-grade California Standards Test (CST) scores and math subject taken, school attendance, and whether students had attended middle school in the same set of feeder middle schools as students in *Green Dot Public Schools*. The authors compared outcomes between the two groups within each cohort and across each year of high school. The analytic sample included 386 students who started ninth grade in 2007–08 (Cohort 1), 1,024 students who started ninth grade in 2008–09 (Cohort 2), and 1,036 students who started ninth grade in 2009–10 (Cohort 3).<sup>7</sup>

## Effectiveness Summary

The WWC review of *Green Dot Public Schools* for the Charter Schools topic area includes outcomes in 10 domains.<sup>8</sup> The one study of *Green Dot Public Schools* that met WWC group design standards reported findings in four domains: mathematics achievement, student progression, school attendance, and English language arts achievement. The following findings present the authors’ estimates and WWC-calculated estimates of the size and statistical significance of the effects of *Green Dot Public Schools* on students. For outcomes measured at multiple points in time for the entire sample, the primary findings that the WWC considered for the effectiveness rating are those measured furthest from the initial exposure to the intervention to reflect the maximum exposure of students to the program. Because we identify the primary finding separately for each sample and outcome measure from among those findings that meet WWC group design standards, the primary findings within an outcome domain are sometimes drawn from different time periods. In cases where outcomes were measured for different portions of the overall sample at different times, each outcome for each sample was considered a primary finding. Additional comparisons are available as supplemental findings in Appendix D. The supplemental findings do not factor into the intervention’s rating of effectiveness. For a more detailed description of the rating of effectiveness and extent of evidence criteria, see the WWC Rating Criteria on p. 24.

### Summary of effectiveness for the mathematics achievement domain

**Table 3. Rating of effectiveness and extent of evidence for the mathematics achievement domain**

Rating of effectiveness	Criteria met
<b>Potentially positive effects</b> <i>Evidence of a positive effect with no overriding contrary evidence.</i>	In the one study that reported findings, the estimated impact of the intervention on outcomes in the <i>mathematics achievement</i> domain was positive and statistically significant.
Extent of evidence	Criteria met
<b>Small</b>	One study that included 2,103 students in 11 schools reported evidence of effectiveness in the <i>mathematics achievement</i> domain.

One study that met WWC group design standards with reservations reported findings in the mathematics achievement domain.

Herman et al. (2012) found a statistically significant positive effect of attending *Green Dot Public Schools* on California Standards Test (CST) scores in Algebra I and Geometry in the first year of high school for students across all three cohorts. The authors also reported statistically significant positive effects of attending *Green Dot Public Schools* on the math scale score of the California High School Exit Exam (CAHSEE) in the second year of high school for students who started ninth grade in 2007–08 (Cohort 1) and 2008–09 (Cohort 2). The study also found statistically significant positive effects of attending *Green Dot Public Schools* on CST scores in Algebra II and Geometry in the second year, and in Algebra II and Summative Math in the third year for students in Cohorts 1 and 2.<sup>9</sup> The WWC applied a correction for multiple comparisons, but this did not affect the statistical significance among any of the contrasts. The WWC characterizes this finding as a potentially positive effect.

Thus, for the mathematics achievement domain, one study showed statistically significant effects. This results in a rating of potentially positive effects, that is based on a small extent of evidence.

Summary of effectiveness for the student progression domain

**Table 4. Rating of effectiveness and extent of evidence for the student progression domain**

Rating of effectiveness	Criteria met
<b>Potentially positive effects</b> <i>Evidence of a positive effect with no overriding contrary evidence.</i>	In the one study that reported findings, the estimated impact of the intervention on outcomes in the <i>student progression</i> domain was statistically significant.
Extent of evidence	Criteria met
<b>Small</b>	One study that included 1,024 students in 11 schools reported evidence of effectiveness in the <i>student progression</i> domain.

One study that met WWC group design standards with reservations reported findings in the student progression domain.

Herman et al. (2012) reported statistically significant positive effects of attending *Green Dot Public Schools* on student progression. Students who started ninth grade in *Green Dot Public Schools* in 2008–09 (Cohort 2) were significantly more likely than comparison students to graduate from high school and to graduate with A-G requirements, the set of high school courses necessary for admissions to California public universities. The WWC applied a correction for multiple comparisons, but this did not affect the statistical significance for either contrast. The WWC characterizes this finding as a potentially positive effect.

Thus, for the student progression domain, one study showed statistically significant effects. This results in a rating of potentially positive effects, that is based on a small extent of evidence.

Summary of effectiveness for the school attendance domain

**Table 5. Rating of effectiveness and extent of evidence for the school attendance domain**

Rating of effectiveness	Criteria met
<b>Potentially positive effects</b> <i>Evidence of a positive effect with no overriding contrary evidence.</i>	In the one study that reported findings, the estimated impact of the intervention on outcomes in the <i>school attendance</i> domain was statistically significant.
Extent of evidence	Criteria met
<b>Small</b>	One study that included 1,578 students in 11 schools reported evidence of effectiveness in the <i>school attendance</i> domain.

One study that met WWC group design standards with reservations reported findings in the school attendance domain.

Herman et al. (2012) reported statistically significant positive effects of attending *Green Dot Public Schools* on school attendance. Students who started at *Green Dot Public Schools* in 2007–08 (Cohort 1) and 2008–09 (Cohort 2) had significantly higher rates of attendance during the fourth year of high school than comparison students. The authors report, and the WWC confirmed, no statistically significant or substantively important effects of *Green Dot Public Schools* on school attendance in Year 1 among students who started ninth grade in 2009–10 (Cohort 3). The WWC characterizes this study as finding a potentially positive effect.

Thus, for the school attendance domain, one study showed statistically significant effects. This results in a rating of potentially positive effects, that is based on a small extent of evidence.

Summary of effectiveness for the English language arts achievement domain

**Table 6. Rating of effectiveness and extent of evidence for the English language arts achievement domain**

Rating of effectiveness	Criteria met
<b>Potentially positive effects</b> <i>Evidence of a positive effect with no overriding contrary evidence.</i>	In the one study that reported findings, the estimated impact of the intervention on outcomes in the <i>English language arts achievement</i> domain was statistically significant.
Extent of evidence	Criteria met
<b>Small</b>	One study that included 1,865 students in 11 schools reported evidence of effectiveness in the <i>English language arts achievement</i> domain.

One study that met WWC group design standards with reservations reported findings in the English language arts domain.

Herman et al. (2012) reported statistically significant positive effects of attending a *Green Dot Public School* on CST scores in English language arts during the first year of high school for students who started ninth grade in *Green Dot Public Schools* in 2009–10 (Cohort 3). The authors reported, and the WWC confirmed, no statistically significant or substantively important effects on CST scores in ELA during the third year of high school for students who started ninth grade in *Green Dot Public Schools* in 2007–08 (Cohort 1) or 2008–09 (Cohort 2). The study also found no statistically significant or substantively important effects of attending *Green Dot Public Schools* on the English Language Arts scale score of the CAHSEE in the second year of high school for Cohorts 1 and 2. The WWC applied a correction for multiple comparisons, but this did not affect the statistical significance among any of the contrasts. The WWC characterizes this study finding as a potentially positive effect.

Thus, for the English language arts achievement domain, one study showed statistically significant effects. This results in a rating of potentially positive effects, that is based on a small extent of evidence.

### References

#### Studies that meet WWC group design standards without reservations

None.

#### Study that meets WWC group design standards with reservations

Herman, J. L., Wang, J., Rickles, J., Hsu, V., Monroe, S., Leon, S., ... CRESST. (2012). *Evaluation of Green Dot's Locke Transformation Project: Findings for Cohort 1 and 2 students* (CRESST Report 815). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED531993>

**Additional sources:**

Herman, J., Rickles, J., Hansen, M., Thomas, L., Gualpa, A., Wang, J., & CRESST. (2011). *Evaluation of Green Dot's Locke Transformation Project: Findings for the 2007-08, 2008-09, and 2009-10 School Years* (CRESST Report 799). National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED522836>

Rickles, J., Wang, J., Herman, J., & CRESST. (2013). *Evaluation of Green Dot's Locke Transformation Project: Supplemental report on Cohort 2 student outcomes* (CRESST Report 825). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED540596>

#### Studies that do not meet WWC group design standards

None.

#### Study that is ineligible for review using the Charter Schools Evidence Review Protocol

Quaile, M. A. (2013). *Reallocating resources to reform schools: A case study of successful school turnarounds in five Los Angeles charter schools* (Doctoral dissertation). Retrieved from <http://digitalibrary.usc.edu/> The study is ineligible for review because it was out of the scope of the protocol.

**Appendix A: Research details for Herman et al. (2012)**

Herman, J. L., Wang, J., Rickles, J., Hsu, V., Monroe, S., Leon, S., ... CRESST. (2012). *Evaluation of Green Dot’s Locke Transformation Project: Findings for Cohort 1 and 2 students* (CRESST Report 815). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED531993>

*Additional sources:*

Herman, J., Rickles, J., Hansen, M., Thomas, L., Gualpa, A., Wang, J., & CRESST. (2011). *Evaluation of Green Dot’s Locke Transformation Project: Findings for the 2007-08, 2008-09, and 2009-10 School Years* (CRESST Report 799). National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED522836>

Rickles, J., Wang, J., Herman, J., & CRESST. (2013). *Evaluation of Green Dot’s Locke Transformation Project: Supplemental report on Cohort 2 student outcomes* (CRESST Report 825). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved from <https://eric.ed.gov/?id=ED540596>

**Table A. Summary of findings**

**Meets WWC Group Design Standards With Reservations**

Outcome domain	Sample size	Study findings	
		Average improvement index (percentile points)	Statistically significant
Mathematics achievement	2,103 students	+13	Yes
Student progression	1,024 students	+10	Yes
School attendance	1,578 students	+5	Yes
English language arts achievement	1,865 students	+4	Yes

**Setting** The study took place in 11 urban high schools in the Los Angeles Unified School District (LAUSD) in Los Angeles, California. Eight schools were *Green Dot Public Schools*, and three were traditional LAUSD high schools.

**Study sample** The study included three cohorts based on year of entry into high school. Cohort 1 entered ninth grade during the 2007–08 school year, Cohort 2 started during the 2008–09 school year, and Cohort 3 started during the 2009–10 school year. Beginning in the fall of 2007, *Green Dot* started converting a single, existing LAUSD high school into *Green Dot Public Schools*. This transition started with two off-site small schools in the first year. By fall 2008, the total neighborhood catchment and full student community that was previously associated with the original LAUSD school became students in eight *Green Dot Public Schools*. The study used a quasi-experimental design to match *Green Dot Public Schools* students to similar students in three traditional high schools who did not have the opportunity to attend the eight *Green Dot Public Schools*. The authors first identified a set of six LAUSD middle schools where the majority of students in *Green Dot Public Schools* had attended eighth grade and the corresponding three traditional high schools where most students in those middle schools started ninth grade. The students in these traditional high schools formed the pool of potential matches for students in *Green Dot Public Schools*. Each student enrolled in *Green Dot Public Schools* was matched exactly to another student from the same cohort enrolled in one of

the traditional high schools based on whether they had the same gender, ethnicity, parents' education, poverty status, language classification, math subject exam taken on the eighth grade California Standards Test (CST), and had attended middle school in the same set of feeder middle schools as students in *Green Dot Public Schools*. When more than one student was identified as a possible match, the authors selected one to be included in the comparison group using nearest-neighbor propensity score matching based on eighth-grade CST scale scores in math and ELA and on eighth-grade attendance rates.

Across the three cohorts, 71%–88% of students were eligible for free or reduced-price lunch, 31%–46% were classified as English language learners, 6%–9% of students had a disability, 74%–79% were Latino, 21%–26% were African American, and 51%–52% were female.

### Intervention group

Students in the intervention condition attended one of eight high schools in the *Green Dot Public Schools* network. The schools in the intervention condition were previously a single traditional LAUSD school that *Green Dot Public Schools* converted into multiple charter schools over a 2-year period. The school was divided into two schools for the first year (the 2007–08 school year), and into eight schools for the second year (the 2008–09 school year).

### Comparison group

Students in the comparison condition attended one of three traditional LAUSD schools.

### Outcomes and measurement

The study measured outcomes at the end of each of 4 years of high school for Cohorts 1 and 2, and the end of the first year of high school for Cohort 3.

Mathematics achievement was measured using state standardized tests. The CST was administered across three math subjects: Algebra I in Year 1, Geometry in Years 1 and 2, Algebra II in Years 2 and 3, and Summative Math in Year 3. Each student took only one of these math subject tests per year, and different groups of students took each subject test every year (for example, some students in a cohort took the Geometry test in Year 1, and others took the test in Year 2). The study also included results from the CAHSEE in math, administered to students for the first time at the end of Year 2. Progression outcomes were the graduation rate and graduation rate after completing a sequence of college preparatory courses within 4 years of entering high school. School attendance was measured in Year 1 for Cohort 3 and in Year 4 for Cohorts 1 and 2. English language arts achievement was measured using state standardized tests. The CST in English language arts was administered in Year 1 for Cohort 3 and in Year 3 for Cohorts 1 and 2. The study also included results from the CAHSEE in English language arts, administered to students for the first time at the end of Year 2. For a more detailed description of these outcome measures, see Appendix B.

Supplemental findings include measurement of outcomes in different years for each cohort, and outcomes based on different samples of students. These additional outcomes are reported as supplemental findings in Appendix D, but do not factor into the intervention's rating of effectiveness.

While the findings from Cohort 2 that contribute to the effectiveness rating are based on a sample with 4 years of outcomes available (results presented in Rickles et al., 2013), findings from Cohort 2 based on an alternate sample with only 3 years of outcomes (results presented in Herman et al., 2012) are reported as supplemental findings.

Results for Algebra II CST scores in Year 2 for Cohort 1 and Geometry CST scores in Year 1 for both Cohort 2 and the alternate Cohort 2 did not meet WWC group design standards. These samples were not shown to be equivalent at baseline across the intervention and comparison groups, and therefore are not included in this review.

The study collected outcomes on course-taking and completion, which were not eligible outcomes for review under the Charter Schools review protocol (version 3.0). The authors also measured semester-to-semester persistence, which was not eligible for review under the protocol because it measures school mobility rather than student progression. In addition to the CAHSEE outcomes described above, the authors measured passing in two or more attempts in both the third and fourth years of high school. The WWC determined these outcomes were not eligible for review because they were conditional on failing the exam on the first attempt. Students who passed the exam on the first attempt or after multiple attempts are included in the passing on any attempt version of the outcome.

### Support for implementation

The *Green Dot Public Schools* model emphasizes partnerships with parents, the school district, and the community. Teachers and principals are provided with recommended best practices to achieve the goals of the organization.

**Appendix B: Outcome measures for each domain**

<b>Mathematics achievement</b>	
<i>California Standards Test (CST): Algebra I (Year 1)</i>	Score on the CST exam in Algebra I administered in the first year of high school (as cited in Rickles et al., 2013). The Algebra I test is one of a set of criterion-referenced standardized tests in the CST administered to California students in grades 2–11 from 2002–14 and aligned to state-adopted content standards.
<i>CST: Geometry (Year 1 or 2)</i>	Score on the CST exam in Geometry, which can be administered in the first or second year of high school (as cited in Rickles et al., 2013).
<i>CST: Algebra II (Year 2 or 3)</i>	Score on the CST exam in Algebra II administered in the second or third year of high school (as cited in Rickles et al., 2013).
<i>CST: Summative Math (Year 3)</i>	Score on the CST exam in Summative Math administered in the third year of high school (as cited in Rickles et al., 2013). The exam is administered to students who have already completed Algebra II and covers content from Algebra I and II, Geometry, and Probability and Statistics.
<i>California High School Exit Exam (CAHSEE): Math scale score (Year 2)</i>	Score on the CAHSEE in math administered in the second year of high school (as cited in Rickles et al., 2013). Beginning with the class of 2006, students in California public schools were required to pass the CAHSEE to demonstrate competency in grade-level skills in reading, writing, and mathematics to earn a high school diploma.
<i>CAHSEE: Ever passed math (Year 3 or 4)</i>	Passing the CAHSEE is a binary measure of whether a student ever passed the math section of the exam, which is measured at the end of both the third and fourth years of high school (as cited in Rickles et al., 2013). The CAHSEE is first administered in the second year of high school, but students can take the test up to three times per school year up to a total of six times. This outcome is only reported as a supplemental finding.
<i>CAHSEE: Passed math on first attempt (Year 2)</i>	Passing the CAHSEE on the first attempt is a binary measure of whether a student passed the math section of the exam on the first attempt during the second year of high school (as cited in Rickles et al., 2013). This outcome is only reported as a supplemental finding.
<b>Student progression</b>	
<i>Graduated (since entry in Year 1 or since Year 4 enrollment)</i>	Graduating from high school is a binary measure of whether a student is observed graduating from any school in the district by the end of the fourth year of high school (as cited in Rickles et al., 2013) for those enrolled in their school at either the beginning of Year 1 or Year 4. All Year 1 students are entering ninth-grade students, and all Year 4 students have stayed enrolled in their same schools from Year 1 to Year 4.
<i>Graduated with A-G (since entry in Year 1 or since Year 4 enrollment)</i>	Graduating from high school with A-G is a binary measure of whether a student graduated and completed a college preparatory sequence of courses by the end of the fourth year of high school (as cited in Rickles et al., 2013) for those enrolled in their school at either the beginning of Year 1 or Year 4. All Year 1 students are entering ninth-grade students, and all Year 4 students have stayed enrolled in their same schools from Year 1 to Year 4.
<b>School attendance</b>	
<i>School attendance (Year 1, 2, 3, or 4)</i>	School attendance is the average of the number of days a student attended divided by the total number of school days for students enrolled in school at the beginning of each year. For each year, the outcome only included students who were still enrolled in the school up to and through that year (as cited in Herman et al., 2012).
<b>English language arts</b>	
<i>CST: English language arts (Year 1, 2, or 3)</i>	Score on the CST exam in the subject of English language arts administered in the first, second, and third years of high school (as cited in Rickles et al., 2013).
<i>CAHSEE: English language arts scale score (Year 2)</i>	Scale score on the CAHSEE in the subject of English language arts administered in the second year of high school (as cited in Rickles et al., 2013).
<i>CAHSEE: Ever passed English language arts (Year 3 or 4)</i>	Passing the CAHSEE is a binary measure of whether a student ever passed the English language arts section of the exam by the third or fourth year of high school (as cited in Rickles et al., 2013). This outcome is only reported as a supplemental finding.
<i>CAHSEE: Passed English language arts on first attempt</i>	Passing the CAHSEE on the first attempt is a binary measure of whether a student passed the English language arts section of the exam on the first attempt during the second year of high school (as cited in Rickles et al., 2013). This outcome is only reported as a supplemental finding.

Appendix C.1: Findings included in the rating for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>CST: Algebra I (Year 1)</i>	Cohorts 1, 2, and 3	11 schools/ 1,718 students	277.86 (48.59)	265.42 (32.93)	12.44	0.30	+12	< .01
<i>CST: Geometry (Year 1)</i>	Cohorts 1, 2, and 3	11 schools/ 385 students	286.84 (43.46)	272.58 (35.60)	14.26	0.37	+14	< .01
<i>CAHSEE: Math Scale Score (Year 2)</i>	Cohorts 1 and 2	11 schools/ 927 students	361.49 (33.13)	354.84 (28.79)	6.65	0.21	+8	< .01
<i>CST: Algebra II (Year 2)</i>	Cohorts 1 and 2	11 schools/ 295 students	270.21 (50.71)	254.16 (41.18)	16.05	0.35	+14	< .01
<i>CST: Geometry (Year 2)</i>	Cohorts 1 and 2	11 schools/ 559 students	255.78 (31.54)	250.09 (29.31)	5.69	0.19	+7	.03
<i>CST: Algebra II (Year 3)</i>	Cohorts 1 and 2	11 schools/ 406 students	253.73 (34.09)	244.65 (30.35)	9.08	0.28	+11	.01
<i>CST: Summative Math (Year 3)</i>	Cohorts 1 and 2	11 schools/ 189 students	271.55 (52.23)	243.95 (42.07)	27.60	0.58	+22	< .01
<b>Domain average for mathematics achievement (Herman et al., 2012)</b>						<b>0.33</b>	<b>+13</b>	<b>Statistically significant</b>
<b>Domain average for mathematics achievement across all studies</b>						<b>0.33</b>	<b>+13</b>	<b>na</b>

**Table Notes:** For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study’s domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable. CST = California Standards Test. CAHSEE = California High School Exit Exam.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient reported in the study) to the unadjusted comparison group posttest means. The findings presented here are based on aggregated samples across cohorts separately reported in the original study. The authors reported p-values for some results, but not for the aggregated analysis. The WWC calculated the p-values reported in the table. When aggregating cohorts, the WWC calculated weighted averages of the adjusted intervention means and unadjusted comparison group means. The authors provided unadjusted outcome standard deviations at the WWC’s request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information. This study is characterized as having a statistically significant positive effect because at least one measure is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix C.2: Findings included in the rating for the student progression domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
% Graduated: Year 1 entry sample	Cohort 2	11 schools/ 1,024 students	54 (na)	40 (na)	14	0.28	+11	< .01
% Graduated with A-G: Year 1 entry sample	Cohort 2	11 schools/ 1,024 students	25 (na)	16 (na)	9	0.22	+9	< .01
<b>Domain average for mathematics achievement (Herman et al., 2012)</b>						<b>0.25</b>	<b>+10</b>	<b>Statistically significant</b>
<b>Domain average for student progression across all studies</b>						<b>0.25</b>	<b>+10</b>	<b>na</b>

**Table Notes:** For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient reported in the study) to the unadjusted comparison group posttest means. The p-values presented here were reported in the original study. The authors provided unadjusted outcome standard deviations at the WWC's request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information. This study is characterized as having a statistically significant positive effect because at least one measure is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix C.3: Findings included in the rating for the school attendance domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>School attendance (Year 1)</i>	Cohort 3	11 schools/ 916 student	0.94 (0.09)	0.94 (0.07)	0.00	0.00	0	.89
<i>School attendance (Year 4)</i>	Cohorts 1 and 2	11 schools/ 662 students	0.94 (0.07)	0.92 (0.08)	0.02	0.27	+10	< .01
<b>Domain average for school attendance (Herman et al., 2012)</b>						<b>0.13</b>	<b>+5</b>	<b>Statistically significant</b>
<b>Domain average for school attendance across all studies</b>						<b>0.13</b>	<b>+5</b>	<b>na</b>

**Table Notes:** For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study’s domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable.

<sup>a</sup> For Herman et al. (2012), the WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient from the authors’ adjusted model) to the unadjusted comparison group posttest means. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient reported in the study) to the unadjusted comparison group posttest means. The findings presented here are based on aggregated samples across cohorts separately reported in the original study. The authors reported p-values for some results, but not for the aggregated analysis. The WWC calculated the p-values reported for the aggregated samples in the table. When aggregating cohorts, the WWC calculated weighted averages of the adjusted intervention means and unadjusted comparison group means. The authors provided unadjusted outcome standard deviations at the WWC’s request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information. This study is characterized as having a statistically significant positive effect because at least one measure is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix C.4: Findings included in the rating for the English language arts achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>CST: ELA (Year 1)</i>	Cohort 3	11 schools/ 930 students	312.00 (48.14)	303.55 (45.46)	8.45	0.18	+7	< .01
<i>CAHSEE: ELA Scale Score (Year 2)</i>	Cohorts 1 and 2	11 schools/ 935 students	355.24 (30.74)	353.46 (29.95)	1.74	0.06	+2	.38
<i>CST: ELA (Year 3)</i>	Cohorts 1 and 2	11 schools/ 752 students	299.98 (51.96)	297.54 (50.05)	2.44	0.05	+2	.51
<b>Domain average for English language arts achievement (Herman et al., 2012)</b>						<b>0.10</b>	<b>+4</b>	<b>Statistically significant</b>
<b>Domain average for English language arts achievement across all studies</b>						<b>0.10</b>	<b>+4</b>	<b>na</b>

**Table Notes:** For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. The WWC-computed average effect size is a simple average rounded to two decimal places; the average improvement index is calculated from the average effect size. The statistical significance of the study's domain average was determined by the WWC. Some statistics may not sum as expected due to rounding. na = not applicable. CST = California Standards Test. CAHSEE = California High School Exit Exam. ELA = English language arts.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient reported in the study) to the unadjusted comparison group posttest means. The findings presented here are based on aggregated samples across cohorts separately reported in the original study. The authors reported p-values for some results, but not for the aggregated analysis. The WWC calculated the p-values reported for the aggregated samples in the table. When aggregating cohorts, the WWC calculated weighted averages of the adjusted intervention means and unadjusted comparison group means. The authors provided unadjusted outcome standard deviations at the WWC's request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information. This study is characterized as having a statistically significant positive effect because at least one measure is positive and statistically significant, and no effects are negative and statistically significant, accounting for multiple comparisons. For more information, please refer to the WWC Procedures and Standards Handbook (version 3.0), p. 26.

Appendix D.1: Description of supplemental findings for the mathematics achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>CST: Algebra I (Year 1)</i>	Cohort 1	5 schools/ 278 students	278.82 (51.53)	264.12 (32.24)	14.70	0.34	+13	< .01
<i>CST: Algebra I (Year 1)</i>	Cohort 2	11 schools/ 714 students	266.70 (35.00)	267.00 (35.00)	-0.30	-0.01	0	.90
<i>CST: Algebra I (Year 1)</i>	Cohort 2 alternate	11 schools/ 795 students	266.88 (36.04)	266.11 (34.05)	0.77	0.02	+1	.74
<i>CST: Algebra I (Year 1)</i>	Cohort 3	11 schools/ 726 students	288.09 (55.12)	264.35 (31.77)	23.74	0.52	+20	< .01
<i>CST: Geometry (Year 1)</i>	Cohort 1	5 schools/ 47 students	300.94 (44.45)	273.83 (49.88)	27.11	0.57	+21	.01
<i>CST: Geometry (Year 1)</i>	Cohort 3	11 schools/ 179 students	280.66 (47.98)	274.32 (33.69)	6.34	0.16	+6	.16
<i>CAHSEE: Math Scale Score (Year 2)</i>	Cohort 1	5 schools/ 228 students	358.62 (33.34)	357.56 (28.04)	1.06	0.03	+1	.75
<i>CAHSEE: Math Scale Score (Year 2)</i>	Cohort 2	11 schools/ 699 students	362.47 (33.00)	354.00 (29.00)	8.47	0.27	+11	< .01
<i>CAHSEE: Math Scale Score (Year 2)</i>	Cohort 2 alternate	11 schools/ 719 students	361.73 (32.86)	354.74 (29.21)	6.99	0.22	+9	< .01
<i>CAHSEE: % Passed math on first attempt (Year 2)</i>	Cohort 1	5 schools/ 228 students	57 (na)	59 (na)	-2	-0.04	-2	.70
<i>CAHSEE: % Passed math on first attempt (Year 2)</i>	Cohort 2	11 schools/ 724 students	59 (na)	51 (na)	8	0.15	+6	.02
<i>CAHSEE: % Passed math on first attempt (Year 2)</i>	Cohort 2 alternate	11 schools/ 726 students	61 (na)	55 (na)	6	0.12	+5	.06
<i>CAHSEE: % Ever passed math (Year 3)</i>	Cohort 1	5 schools/ 187 students	84 (na)	70 (na)	14	0.33	+13	.02
<i>CAHSEE: % Ever passed math (Year 3)</i>	Cohort 2 alternate	11 schools/ 616 students	82 (na)	73 (na)	9	0.22	+9	< .01
<i>CST: Algebra II (Year 2)</i>	Cohort 2	11 schools/ 229 students	267.99 (48.00)	257.00 (42.00)	10.99	0.24	+10	.02
<i>CST: Algebra II (Year 2)</i>	Cohort 2 alternate	11 schools/ 247 students	268.16 (47.34)	256.79 (41.49)	11.37	0.25	+10	.02
<i>CST: Geometry (Year 2)</i>	Cohort 1	5 schools/ 151 students	257.38 (32.98)	244.31 (22.69)	13.07	0.45	+17	.01
<i>CST: Geometry (Year 2)</i>	Cohort 2	11 schools/ 408 students	255.12 (31.00)	252.00 (31.00)	3.12	0.10	+4	.26
<i>CST: Geometry (Year 2)</i>	Cohort 2 alternate	11 schools/ 446 students	255.23 (30.51)	251.93 (29.51)	3.30	0.11	+4	.22
<i>CST: Algebra II (Year 3)</i>	Cohort 1	5 schools/ 106 students	247.30 (27.47)	243.64 (25.38)	3.66	0.14	+5	.48

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<i>CST: Algebra II (Year 3)</i>	Cohort 2 alternate	11 schools/ 316 students	255.34 (35.89)	242.90 (31.83)	12.44	0.36	+14	< .01
<i>CST: Summative Math (Year 3)</i>	Cohort 1	5 schools/ 50 students	273.67 (63.43)	243.83 (52.60)	29.84	0.50	+19	.04
<i>CST: Summative Math (Year 3)</i>	Cohort 2	11 schools/ 139 students	270.80 (48.00)	244.00 (38.00)	26.80	0.61	+23	< .01
<i>CST: Summative Math (Year 3)</i>	Cohort 2 alternate	11 schools/ 151 students	269.89 (46.77)	247.65 (43.38)	22.24	0.49	+19	< .01
<i>CAHSEE: % Ever passed math (Year 4)</i>	Cohort 1	5 schools/ 166 students	90 (na)	92 (na)	-2	-0.07	-3	.44
<i>CAHSEE: % Ever passed math (Year 4)</i>	Cohort 2	11 schools/ 496 students	96 (na)	91 (na)	5	0.20	+8	.03
<i>CAHSEE: % Ever passed math (Year 1 entry sample)</i>	Cohort 2	11 schools/ 1,024 students	69 (na)	61 (na)	8	0.17	+7	.01

**Table Notes:** The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not applicable. CST = California Standards Test. CAHSEE = California High School Exit Exam.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed and resulted in a change in the statistical significance for one finding: the author-reported *p*-value for CST: Summative Math (Year 3) for Cohort 1 was .041, and the WWC-computed critical *p*-value was .032. Therefore, the WWC does not find the result for this finding to be statistically significant. The *p*-values presented here were reported in the original study. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient from the authors’ adjusted model) to the unadjusted comparison group posttest means. The authors provided unadjusted outcome standard deviations at the WWC’s request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information.

Appendix D.2: Description of supplemental findings for the school progression domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
% Graduated: Year 4 Sample	Cohort 1	5 schools/ 166 students	79 (na)	55 (na)	24	0.52	+20	< .01
% Graduated: Year 4 Sample	Cohort 2	11 schools/ 496 students	88 (na)	70 (na)	18	0.44	+17	< .01
% Graduated w/ A-G: Year 4 Sample	Cohort 1	5 schools/ 166 students	47 (na)	13 (na)	34	0.78	+28	< .01
% Graduated w/ A-G: Year 4 Sample	Cohort 2	11 schools/ 496 students	44 (na)	27 (na)	17	0.36	+14	< .01

**Table Notes:** The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual’s percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not applicable.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed but did not affect whether any of the contrasts were found to be statistically significant. The p-values presented here were reported in the original study. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient from the authors’ adjusted model) to the unadjusted comparison group posttest. The authors provided unadjusted outcome standard deviations at the WWC’s request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information.

Appendix D.3: Description of supplemental findings for the school attendance domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>School attendance: Year 1</i>	Cohort 1	5 schools/ 330 students	0.93 (0.10)	0.93 (0.07)	0.00	0.00	0	.99
<i>School attendance: Year 1</i>	Cohort 2	11 schools/ 888 students	0.93 (0.08)	0.93 (0.08)	0.00	0.00	0	.66
<i>School attendance: Year 1</i>	Cohort 2 alternate	11 schools/ 978 students	0.92 (0.08)	0.92 (0.08)	0.00	0.00	0	.90
<i>School attendance: Year 2</i>	Cohort 1	5 schools/ 242 students	0.93 (0.13)	0.94 (0.05)	-0.01	-0.10	-4	.43
<i>School attendance: Year 2</i>	Cohort 2	11 schools/ 718 students	0.93 (nr)	0.93 (nr)	0.00	na	na	.87
<i>School attendance: Year 2</i>	Cohort 2 alternate	11 schools/ 779 students	0.93 (0.11)	0.93 (0.08)	0.00	0.00	0	.54
<i>School attendance: Year 3</i>	Cohort 1	5 schools/ 187 students	0.95 (0.06)	0.93 (0.07)	0.02	0.31	+12	.04
<i>School attendance: Year 3</i>	Cohort 2	11 schools/ 564 students	0.94 (0.07)	0.94 (0.07)	0.00	0.00	0	.94
<i>School attendance: Year 3</i>	Cohort 2 alternate	11 schools/ 622 students	0.93 (0.08)	0.93 (0.08)	0.00	0.00	0	.48
<i>School attendance: Year 4</i>	Cohort 1	5 schools/ 166 students	0.94 (0.08)	0.91 (0.09)	0.03	0.35	+14	.04
<i>School attendance: Year 4</i>	Cohort 2	11 schools/ 496 students	0.94 (0.07)	0.93 (0.08)	0.01	0.13	+5	.02

**Table Notes:** The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. nr = not reported.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed and resulted in a change in the statistical significance for three findings: the author-reported *p*-value for Student attendance: Year 3 for Cohort 1 was .043, and the WWC-computed critical *p*-value was .014; the author-reported *p*-value for Student attendance: Year 4 for Cohort 1 was .036, and the WWC-computed critical *p*-value was .009; the author-reported *p*-value for Student attendance: Year 4 for Cohort 2 was .024, and the WWC-computed critical *p*-value was .005. Therefore, the WWC does not find the result for any of these three findings to be statistically significant. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient from the authors' adjusted model) to the unadjusted comparison group posttest. The authors provided unadjusted outcome standard deviations at the WWC's request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information. The authors did not provide standard deviations for the Student attendance: Year 2 outcome for Cohort 2, so the WWC cannot calculate an effect size.

Appendix D.4: Description of supplemental findings for the English language arts achievement domain

Outcome measure	Study sample	Sample size	Mean (standard deviation)		WWC calculations			p-value
			Intervention group	Comparison group	Mean difference	Effect size	Improvement index	
<b>Herman et al., 2012<sup>a</sup></b>								
<i>CST: ELA (Year 1)</i>	Cohort 1	5 schools/ 330 students	314.72 (41.50)	305.50 (47.40)	9.22	0.21	+8	< .01
<i>CST: ELA (Year 1)</i>	Cohort 2	11 schools/ 888 students	304.42 (44.00)	301.00 (44.00)	3.42	0.08	+3	.07
<i>CST: ELA (Year 1)</i>	Cohort 2 alternate	11 schools/ 978 students	304.50 (43.92)	301.57 (43.74)	2.93	0.07	+3	.10
<i>CAHSEE: ELA Scale Score (Year 2)</i>	Cohort 1	5 schools/ 230 students	353.00 (29.77)	354.94 (29.86)	-1.94	-0.07	-3	.47
<i>CAHSEE: ELA Scale Score (Year 2)</i>	Cohort 2	11 schools/ 705 students	355.94 (31.00)	353.00 (30.00)	2.94	0.10	+4	.06
<i>CAHSEE: ELA Scale Score (Year 2)</i>	Cohort 2 alternate	11 schools/ 725 students	357.53 (30.23)	353.74 (31.01)	3.79	0.12	+5	.01
<i>CAHSEE: % Passed ELA on first attempt (Year 2)</i>	Cohort 1	5 schools/ 230 students	54 (na)	62 (na)	-8	-0.16	-6	.16
<i>CAHSEE: % Passed ELA on first attempt (Year 2)</i>	Cohort 2	11 schools/ 724 students	59 (na)	51 (na)	8	0.15	+6	.01
<i>CAHSEE: % Passed ELA on first attempt (Year 2)</i>	Cohort 2 alternate	11 schools/ 725 students	60 (na)	54 (na)	6	0.12	+5	.04
<i>CST: ELA (Year 2)</i>	Cohort 1	5 schools/ 242 students	303.29 (46.74)	300.01 (47.25)	3.28	0.07	+3	.44
<i>CST: ELA (Year 2)</i>	Cohort 2	11 schools/ 724 students	299.82 (47.00)	296.00 (43.00)	3.82	0.09	+3	.10
<i>CST: ELA (Year 2)</i>	Cohort 2 alternate	11 schools/ 786 students	299.13 (46.76)	295.14 (43.80)	3.99	0.09	+4	.07
<i>CST: ELA (Year 3)</i>	Cohort 1	5 schools/ 188 students	291.77 (51.35)	296.15 (50.43)	-4.38	-0.09	-3	.44
<i>CST: ELA (Year 3)</i>	Cohort 2	11 schools/ 564 students	302.71 (52.00)	298.00 (50.00)	4.71	0.09	+4	.11
<i>CST: ELA (Year 3)</i>	Cohort 2 alternate	11 schools/ 622 students	301.52 (52.06)	298.48 (49.44)	3.04	0.06	+2	.40
<i>CAHSEE: % Ever passed ELA (Year 3)</i>	Cohort 1	5 schools/ 187 students	76 (na)	73 (na)	3	0.07	+3	.59
<i>CAHSEE: % Ever passed ELA (Year 3)<sup>b</sup></i>	Cohort 2 alternate	11 schools/ 616 students	79 (na)	73 (na)	6	0.14	+6	.01
<i>CAHSEE: % Ever passed ELA (Year 4)</i>	Cohort 1	5 schools/ 166 students	86 (na)	90 (na)	-4	-0.12	-5	.35
<i>CAHSEE: % Ever passed ELA (Year 4)</i>	Cohort 2	11 schools/ 496 students	94 (na)	90 (na)	4	0.14	+6	.07
<i>CAHSEE: % Ever passed ELA (Year 1 entry sample)</i>	Cohort 2	11 schools/ 1,024 students	69 (na)	61 (na)	8	0.17	+7	< .01

**Table Notes:** The supplemental findings presented in this table are additional findings from studies in this report that meet WWC design standards with or without reservations, but do not factor into the determination of the intervention rating. For mean difference, effect size, and improvement index values reported in the table, a positive number favors the intervention group and a negative number favors the comparison group. The effect size is a standardized measure of the effect of an intervention on outcomes, representing the average change expected for all individuals who are given the intervention (measured in standard deviations of the outcome measure). The improvement index is an alternate presentation of the effect size, reflecting the change in an average individual's percentile rank that can be expected if the individual is given the intervention. Some statistics may not sum as expected due to rounding. na = not applicable. CST = California Standards Test. CAHSEE = California High School Exit Exam. ELA = English language arts.

<sup>a</sup> For Herman et al. (2012), a correction for multiple comparisons was needed and resulted in a change in the statistical significance for four findings: the author-reported  $p$ -value for CAHSEE: ELA scale score (Year 2) for the alternate Cohort 2 outcome was .014, and the WWC-computed critical  $p$ -value was .013; the author-reported  $p$ -value for the CAHSEE: % Passed ELA on the first attempt (Year 2) for Cohort 2 outcome was .012, and the WWC-computed critical  $p$ -value was .010; the author-reported  $p$ -value for the CAHSEE: % Passed ELA on the first attempt (Year 2) for the alternate Cohort 2 outcome was .035, and the WWC-computed critical  $p$ -value was .015; the author-reported  $p$ -value for the CAHSEE: % Ever Passed ELA (Year 3) for the alternate Cohort 2 outcome was .008, and the WWC-computed critical  $p$ -value was .0075. Therefore, the WWC does not find the result for any of these four outcomes to be statistically significant. The  $p$ -values presented here were reported in the original study. The WWC calculated adjusted intervention group means by adding the impact of the intervention (the regression coefficient from the authors' adjusted model) to the unadjusted comparison group posttest. The authors provided unadjusted outcome standard deviations at the WWC's request. Please see the WWC Procedures and Standards Handbook (version 3.0) p. 23 for more information.

<sup>b</sup> In response to an author query for Herman et al. (2012), the authors provided a corrected sample size different from what was originally published. The study originally reported the intervention sample size was 308 for this outcome, but updated the sample to 309 in their response to the query. The updated number is included here.

### Endnotes

<sup>1</sup> The descriptive information for this intervention comes from publicly available sources: the intervention website (*Green Dot Public Schools* website at <http://greendot.org/about/>, downloaded in April 2017) and Dillon, S. (2010, June). *School is turned around, but cost gives pause*. Retrieved from <http://www.nytimes.com/>). The What Works Clearinghouse (WWC) requests developers review the intervention description sections for accuracy from their perspective. The WWC provided the developer with the intervention description in April 2017; however, the WWC did not receive a response. Further verification of the accuracy of the descriptive information for this intervention is beyond the scope of this review.

<sup>2</sup> The literature search reflects documents publicly available by February 2017. Reviews of the studies in this report used the standards from the WWC Procedures and Standards Handbook (version 3.0) and the Charter Schools review protocol (version 3.0). The evidence presented in this report is based on available research. Findings and conclusions could change as new research becomes available.

<sup>3</sup> Please see the Charter Schools review protocol (version 3.0) for a list of all outcome domains.

<sup>4</sup> For criteria used to determine the rating of effectiveness and extent of evidence, see the WWC Rating Criteria on p. 24. These improvement index numbers show the average and range of individual-level improvement indices for all findings across the studies.

<sup>5</sup> The sample sizes reported here represent the largest non-overlapping samples of students from Cohorts 1, 2, and 3 in each domain across findings that contribute to the intervention rating (i.e., those reported in Appendix C).

<sup>6</sup> Please see the Charter Schools review protocol (version 3.0) for details on the types of interventions that are eligible for review. A study of the effectiveness of an individual charter school is eligible to be included in a review of the evidence of the effectiveness of an individual charter school, but is not eligible to be included in a review of the evidence of the effectiveness of a named CMO or charter network, like *Green Dot Public Schools*.

<sup>7</sup> The findings from Cohort 2 that contribute to the effectiveness rating are based on a sample with 4 years of outcomes available (results presented in Rickles et al., 2013). Findings from Cohort 2 based on an alternate sample with only 3 years of outcomes (results presented in Herman et al., 2012) are reported as supplemental findings, which do not contribute to the effectiveness rating in this intervention report. These alternate findings for Cohort 2 are reported in Appendix D.

<sup>8</sup> The 10 student outcome domains are: mathematics achievement, English language arts achievement, science achievement, social studies achievement, general achievement, social-emotional competence, disciplinary experiences, student attendance, student progression, and earnings in adulthood.

<sup>9</sup> Students take mathematics CST exams in different subjects each year. For example, students who take the Algebra I exam in Year 1 generally take the Geometry exam in Year 2, while other students take Geometry in Year 1 and Algebra II in Year 2. Because the set of students that take each exam differs by year, the subject-specific tests contribute to the effectiveness rating in each year they are given.

### Recommended Citation

What Works Clearinghouse, Institute of Education Sciences, U.S. Department of Education. (2018, January). *Charter Schools intervention report: Green Dot Public Schools*. Retrieved from <http://whatworks.ed.gov>

## WWC Rating Criteria

### Criteria used to determine the rating of a study

Study rating	Criteria
<b>Meets WWC group design standards without reservations</b>	A study that provides strong evidence for an intervention's effectiveness, such as a well-implemented RCT.
<b>Meets WWC group design standards with reservations</b>	A study that provides weaker evidence for an intervention's effectiveness, such as a QED or an RCT with high attrition that has established equivalence of the analytic samples.

### Criteria used to determine the rating of effectiveness for an intervention

Rating of effectiveness	Criteria
<b>Positive effects</b>	Two or more studies show statistically significant positive effects, at least one of which met WWC group design standards without reservations, AND No studies show statistically significant or substantively important negative effects.
<b>Potentially positive effects</b>	At least one study shows a statistically significant or substantively important positive effect, AND No studies show a statistically significant or substantively important negative effect AND fewer or the same number of studies show indeterminate effects than show statistically significant or substantively important positive effects.
<b>Mixed effects</b>	At least one study shows a statistically significant or substantively important positive effect AND at least one study shows a statistically significant or substantively important negative effect, but no more such studies than the number showing a statistically significant or substantively important positive effect, OR At least one study shows a statistically significant or substantively important effect AND more studies show an indeterminate effect than show a statistically significant or substantively important effect.
<b>Potentially negative effects</b>	One study shows a statistically significant or substantively important negative effect and no studies show a statistically significant or substantively important positive effect, OR Two or more studies show statistically significant or substantively important negative effects, at least one study shows a statistically significant or substantively important positive effect, and more studies show statistically significant or substantively important negative effects than show statistically significant or substantively important positive effects.
<b>Negative effects</b>	Two or more studies show statistically significant negative effects, at least one of which met WWC group design standards without reservations, AND No studies show statistically significant or substantively important positive effects.
<b>No discernible effects</b>	None of the studies shows a statistically significant or substantively important effect, either positive or negative.

### Criteria used to determine the extent of evidence for an intervention

Extent of evidence	Criteria
<b>Medium to large</b>	The domain includes more than one study, AND The domain includes more than one school, AND The domain findings are based on a total sample size of at least 350 students, OR, assuming 25 students in a class, a total of at least 14 classrooms across studies.
<b>Small</b>	The domain includes only one study, OR The domain includes only one school, OR The domain findings are based on a total sample size of fewer than 350 students, AND, assuming 25 students in a class, a total of fewer than 14 classrooms across studies.

### Glossary of Terms

**Attrition** Attrition occurs when an outcome variable is not available for all subjects initially assigned to the intervention and comparison groups. If a randomized controlled trial (RCT) or regression discontinuity design (RDD) study has high levels of attrition, the validity of the study results can be called into question. An RCT with high attrition cannot receive the highest rating of *Meets WWC Group Design Standards without Reservations*, but can receive a rating of *Meets WWC Group Design Standards with Reservations* if it establishes baseline equivalence of the analytic sample. Similarly, the highest rating an RDD with high attrition can receive is *Meets WWC RDD Standards with Reservations*.

For single-case design research, attrition occurs when an individual fails to complete all required phases or data points in an experiment, or when the case is a group and individuals leave the group. If a single-case design does not meet minimum requirements for phases and data points within phases, the study cannot receive the highest rating of *Meets WWC Pilot Single-Case Design Standards without Reservations*.

**Baseline** A point in time before the intervention was implemented in group design research and in regression discontinuity design studies. When a study is required to satisfy the baseline equivalence requirement, it must be done with characteristics of the analytic sample at baseline. In a single-case design experiment, the baseline condition is a period during which participants are not receiving the intervention.

**Clustering Adjustment** An adjustment to the statistical significance of a finding when the units of assignment and analysis differ. When random assignment is carried out at the cluster level, outcomes for individual units within the same clusters may be correlated. When the analysis is conducted at the individual level rather than the cluster level, there is a mismatch between the unit of assignment and the unit of analysis, and this correlation must be accounted for when assessing the statistical significance of an impact estimate. If the correlation is not accounted for in a mismatched analysis, the study may be too likely to report statistically significant findings. To fairly assess an intervention's effects, in cases where study authors have not corrected for the clustering, the WWC applies an adjustment for clustering when reporting statistical significance.

**Confounding factor** A confounding factor is a component of a study that is completely aligned with one of the study conditions, making it impossible to separate how much of the observed effect was due to the intervention and how much was due to the factor.

**Design** The method by which intervention and comparison groups are assigned (group design and regression discontinuity design) or the method by which an outcome measure is assessed repeatedly within and across different phases that are defined by the presence or absence of an intervention (single-case design). Designs eligible for WWC review are randomized controlled trials, quasi-experimental designs, regression discontinuity designs, and single-case designs.

**Effect size** The effect size is a measure of the magnitude of an effect. The WWC uses a standardized measure to facilitate comparisons across studies and outcomes.

**Eligibility** A study is eligible for review and inclusion in this report if it falls within the scope of the review protocol and uses either an experimental or matched comparison group design.

**Equivalence** A demonstration that the analytic sample groups are similar on observed characteristics defined in the review area protocol.

### Glossary of Terms

**Extent of evidence** An indication of how much evidence from group design studies supports the findings in an intervention report. The extent of evidence categorization for intervention reports focuses on the number and sizes of studies of the intervention in order to give an indication of how broadly findings may be applied to different settings. There are two extent of evidence categories: small and medium to large.

- **small:** includes only one study, or one school, or findings based on a total sample size of less than 350 students and 14 classrooms (assuming 25 students in a class)
- **medium to large:** includes more than one study, more than one school, and findings based on a total sample of at least 350 students or 14 classrooms

**Gain scores** The result of subtracting the pretest from the posttest for each individual in the sample. Some studies analyze gain scores instead of the unadjusted outcome measure as a method of accounting for the baseline measure when estimating the effect of an intervention. The WWC reviews and reports findings from analyses of gain scores, but gain scores do not satisfy the WWC's requirement for a statistical adjustment under the baseline equivalence requirement. This means that a study that must satisfy the baseline equivalence requirement and has baseline differences between 0.05 and 0.25 standard deviations *Does Not Meet WWC Group Design Standards* if the study's only adjustment for the baseline measure was in the construction of the gain score.

**Group design** A study design in which outcomes for a group receiving an intervention are compared to those for a group not receiving the intervention. Comparison group designs eligible for WWC review are randomized controlled trials and quasi-experimental designs.

**Improvement index** Along a percentile distribution of individuals, the improvement index represents the gain or loss of the average individual due to the intervention. As the average individual starts at the 50th percentile, the measure ranges from -50 to +50.

**Intervention** An educational program, product, practice, or policy aimed at improving student outcomes.

**Intervention report** A summary of the findings of the highest-quality research on a given program, product, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against design standards, and summarizes the findings of those that meet WWC design standards.

**Multiple comparison adjustment** An adjustment to the statistical significance of results to account for multiple comparisons in a group design study. The WWC uses the Benjamini-Hochberg (BH) correction to adjust the statistical significance of results within an outcome domain when study authors perform multiple hypothesis tests without adjusting the  $p$ -value. The BH correction is used in three types of situations: studies that tested multiple outcome measures in the same outcome domain with a single comparison group; studies that tested a given outcome measure with multiple comparison groups; and studies that tested multiple outcome measures in the same outcome domain with multiple comparison groups. Because repeated tests of highly correlated constructs will lead to a greater likelihood of mistakenly concluding that the impact was different from zero, in all three situations, the WWC uses the BH correction to reduce the possibility of making this error. The WWC makes separate adjustments for primary and secondary findings.

Please see the [WWC Procedures and Standards Handbook \(version 3.0\)](#) for additional details.

### Glossary of Terms

<b>Outcome domain</b>	A group of closely-related outcomes. A domain is the organizing construct for a set of related outcomes through which studies claim effectiveness.
<b>Quasi-experimental design (QED)</b>	A quasi-experimental design (QED) is a research design in which study participants are assigned to intervention and comparison groups through a process that is not random.
<b>Randomized controlled trial (RCT)</b>	A randomized controlled trial (RCT) is an experiment in which eligible study participants are randomly assigned to intervention and comparison groups.
<b>Rating of effectiveness</b>	For group design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the magnitude, statistical significance, and consistency in findings. For single-case design research, the WWC rates the effectiveness of an intervention in each domain based on the quality of the research design and the consistency of demonstrated effects. The criteria for the ratings of effectiveness are given in the WWC Rating Criteria on p. 24.
<b>Regression discontinuity design (RDD)</b>	A design in which groups are created using a continuous scoring rule. For example, students may be assigned to a summer school program if they score below a preset point on a standardized test, or schools may be awarded a grant based on their score on an application. A regression line or curve is estimated for the intervention group and similarly for the comparison group, and an effect occurs if there is a discontinuity in the two regression lines at the cutoff.
<b>Single-case design</b>	A research approach in which an outcome variable is measured repeatedly within and across different conditions that are defined by the presence or absence of an intervention.
<b>Standard deviation</b>	The standard deviation of a measure shows how much variation exists across observations in the sample. A low standard deviation indicates that the observations in the sample tend to be very close to the mean; a high standard deviation indicates that the observations in the sample tend to be spread out over a large range of values.
<b>Statistical significance</b>	Statistical significance is the probability that the difference between groups is a result of chance rather than a real difference between the groups. The WWC labels a finding statistically significant if the likelihood that the difference is due to chance is less than 5% ( $p < .05$ ).
<b>Study rating</b>	The result of the WWC assessment of a study. The rating is based on the strength of the evidence of the effectiveness of the educational intervention. Studies are given a rating of <i>Meets WWC Design Standards without Reservations</i> , <i>Meets WWC Design Standards with Reservations</i> , or <i>Does Not Meet WWC Design Standards</i> , based on the assessment of the study against the appropriate design standards. The WWC has design standards for group design, single-case design, and regression discontinuity design studies.
<b>Substantively important</b>	A substantively important finding is one that has an effect size of 0.25 or greater, regardless of statistical significance.
<b>Systematic review</b>	A review of existing literature on a topic that is identified and reviewed using explicit methods. A WWC systematic review has five steps: 1) developing a review protocol; 2) searching the literature; 3) reviewing studies, including screening studies for eligibility, reviewing the methodological quality of each study, and reporting on high quality studies and their findings; 4) combining findings within and across studies; and, 5) summarizing the review.

Please see the [WWC Procedures and Standards Handbook \(version 3.0\)](#) for additional details.



An **intervention report** summarizes the findings of high-quality research on a given program, practice, or policy in education. The WWC searches for all research studies on an intervention, reviews each against evidence standards, and summarizes the findings of those that meet standards.

This intervention report was prepared for the WWC by Mathematica Policy Research under contract ED-IES-13-C-0010.