The Study of Deeper Learning: College Enrollment, Persistence, and Degree Completion in the First 6 Years After High School

Prepared for

The William and Flora Hewlett Foundation

Prepared by

American Institutes for Research

1400 Crystal Drive, 10th Floor

Arlington, VA 22202

https://www.air.org

JUNE 2022

Advancing Evidence. Improving Lives.

The Study of Deeper Learning: College Enrollment, Persistence, and Degree Completion in the First 6 Years After High School

Appendix Authors

Kristina L. Zeiser and Ji Hyun Yang

Principal Investigators

Jennifer O'Day and Michael S. Garet

Study Team

Catherine Bitter, Candice Handjojo, Dana Shaat, and Angela Whistler

JUNE 2022

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**

18141_06/22

Contents

Exhibits

I. Introduction

The *Study of Deeper Learning: Opportunities and Outcomes* is a proof-of-concept study focused on students who attended high schools with at least moderately well-implemented network approaches targeting deeper learning (network schools) and schools not implementing network approaches targeting deeper learning but serving similar populations of students (non-network schools). The study was conducted in pairs of network and non-network schools that serve similar student populations in several districts in California and New York City. Relying on follow-up survey and interview data and data from the StudentTracker service at the National Student Clearinghouse (NSC), a follow-up study conducted between 2019 and 2022 examined differences in students' college, workforce, and civic engagement outcomes up to 6 years after expected high school graduation.

This technical appendix provides additional study information for Report 7, *Deeper Learning and College Enrollment, Persistence, and Degree Completion in the First 6 Years After High School*. It provides an extended description of the study's sampling procedures, data sources, analytic methods, and results. It begins by describing how network and non-network schools were selected and recruited to participate in the study. After presenting the characteristics of the participating schools, we describe the student samples, the selection of student samples for primary data collection, and the levels of student attrition between Grade 9 entry and data collection. After describing the high school survey instrument and NSC data, we provide information about the creation of weights and the statistical models used within the report. The appendix concludes with tables that contain the findings discussed in the report.

II. Study Sample

A. Network School Recruitment and Comparison School Selection

In 2011–12, the Hewlett Foundation selected 10 school networks to participate in what would become the "Deeper Learning Community of Practice." The purpose of this community of practice is to share strategies, tools, and lessons that both contribute to the work of the networks themselves and build the broader knowledge base about deeper learning. The main selection criteria for the networks were as follows:

- The networks needed to have experience in—and an explicit focus on—promoting a deep understanding of content and the kinds of competencies reflected in the Hewlett Foundation's identified dimensions of deeper learning.

- They needed to do this across whole schools serving diverse populations of students (rather than targeting only certain portions of the students or teachers in a school).

The Hewlett Foundation selected the Community of Practice networks prior to the start of the *Study of Deeper Learning: Opportunities and Outcomes*. The 10 networks represented in this study have a well-established history of promoting deeper learning, and all share an emphasis on providing educational opportunities for minority students and students from low-income families to prepare them for college and career. For the original study, the American Institutes for Research (AIR) recruited a set of 20 network high schools from the 10 networks. The criteria for network school selection are in Exhibit 2.1.

Exhibit 2.1. Network and Non-Network School Eligibility Requirements

	Network school criteria	Non-network school criteria
Regular high school (i.e., not a special education, vocational, or alternative high school)	✓	✓
Nonmagnet school	✓	✓
Noncharter school		✓
Low grade is Grade 9		✓
Low grade is Grades K–9	✓	
High grade is Grade 12	✓	✓
More than 25% of the students are eligible for free or reduced-price lunch	✓	✓
More than 200 students enrolled in Grades 9–12	✓	✓
Been in the network since the 2007–08 school year	✓	
Schoolwide implementation of the network approach	✓	
A moderate or high implementation rating from the network	✓	
Within the same district as a network school or a surrounding district		✓

Note. Some deeper learning networks begin focusing on deeper learning competencies before Grade 9. Although these network schools included grades below Grade 9, we selected for our study students who did not attend a deeper learning network school until Grade 9. No non-network schools selected for the study had students below Grade 9.

Given both the small number of network schools in the sample and the criteria used to select the sample, the study's findings are limited in terms of their generalizability. For example, the 10 networks include many schools that were excluded by the study's criteria (such as elementary and middle schools, very small schools, schools without substantial disadvantaged populations, and schools that opened very recently). Furthermore, because AIR included only moderate to high implementers of the network models, findings cannot be generalized to all schools trying to implement a deeper learning approach.

To select non-network schools, we first identified schools with a population of incoming Grade 9 students similar to the incoming Grade 9 students at the network schools. We identified a set of eligible non-network schools located in the same school district as the network school (if the network school was operated by a school district) or within the surrounding school district of the network school (if the network school was operated by a charter school management organization). Schools were identified using the 2007–08, 2008–09, and 2009–10 Common Core of Data (CCD) and were deemed eligible if they met the criteria in Exhibit 2.1. Specifically, the study team used the 2007–08 data to determine whether the school existed as of the 2007–08 school year, and we used averages from the 2008–09 and 2009–10 school years to determine the overall number of students and the percentage of students eligible for free or reduced-price lunch. We expected the distribution of students across racial/ethnic categories to be relatively stable across years for most schools, so we relied on the 2009–10 data.[1]

Based on the CCD data, the study team identified up to five matches for each network school, relying on Mahalanobis distances that were computed using four variables: the average percentage of students eligible for free or reduced-price lunch, the percentage of African American students, the percentage of Hispanic students, and the percentage of White students from the 2008–09 and 2009–10 CCD. To guard against matching dissimilar schools, we required comparison schools to be within one standard deviation of its paired network school on all four variables we used to calculate Mahalanobis distance. After receiving extant district data, we compared the Grade 8 achievement of students in the network school and students in the selected comparison schools to determine priorities for school recruitment.

[1] Although we expected school characteristics to be reasonably stable from 2007–08 to 2009–10, schools that had recently opened might have experienced changes in enrollment during the first few years after opening. For example, if a school opened in 2007–08, and it first enrolled only Grade 9 students and added a grade each year, its highest grade would have been Grade 9 in 2007–08, Grade 10 in 2008–09, and Grade 11 in 2009–10. Similarly, the school’s enrollment would have increased during the same period. As such, selection criteria were modified for recently opened schools. To ensure a sufficient sample size for schools that had recently opened, we removed schools with fewer than 200 students, on average, between the 2008–09 and 2009–10 school years (rather than within each school year), even if the school only had two and three cohorts of students in those years, respectively.

An overview of the matched school pairs that were included in the *Study of Deeper Learning* is in Exhibit 2.2.[2] Pairs 3, 5, and 12 were excluded from Report 7 because we could not obtain college outcome data for students in these school pairs.

Exhibit 2.2. Description of School Pairs

		Enrollment	% Female	% African American	% Hispanic	% Asian	% FRPL
Pair 1 (CA)	Network (1N)	400	70	30	40	10	70
	Non-network (1C)	2,100	50	20	20	30	40
Pair 2 (CA)	Network (2N)	300	50	10	40	0	40
	Non-network (2C)	1,600	50	20	30	10	50
Pair 3 (CA)	Network (3N)[a]	400	50	20	50	10	60
	Non-network (3C)	1,800	50	40	20	20	50
Pair 4 (CA)	Network (4N)	300	50	0	90	10	50
	Non-network (4C)	2,300	50	0	90	10	70
Pair 5 (CA)	Network (5N)	400	50	0	100	0	40
	Non-network (4C)	2,300	50	0	90	10	70
Pair 6 (CA)	Network (6N)	600	50	10	10	10	30
	Non-network (6C)	2,600	50	10	30	0	20
Pair 7 (CA)	Network (7N1)	400	50	10	10	10	40
	Network (7N2)	400	50	10	10	10	40
	Non-network (7C)	2,500	50	10	30	10	50
Pair 8 (NY)	Network (8N)	500	60	10	20	10	40
	Non-network (8C)	600	60	10	20	20	50
Pair 9 (NY)	Network (9N)	400	60	40	60	0	80
	Non-network (9C)	400	40	40	50	0	70
Pair 10 (NY)	Network (10N)	400	40	0	40	60	100
	Non-network (10C1)	600	50	0	100	0	80
	Non-network (10C2)	500	50	0	90	10	90
Pair 11 (NY)	Network (11N)	400	50	20	40	30	100
	Non-network (10C1)	600	50	0	100	0	80
	Non-network (10C2)	500	50	0	90	10	90

[2] In addition to the school pairs listed in Exhibit 2.2, qualitative data were collected from four network schools for which we could not identify an appropriate matched non-network school because of either unique school features or an inability to access administrative, student-level data.

		Enrollment	% Female	% African American	% Hispanic	% Asian	% FRPL
Pair 12 (CA)	Network (12N)	300	50	60	30	0	40
	Non-network (3C)	1,800	50	40	20	20	50
Pair 13 (NY)	Network (13N)	400	60	80	20	0	80
	Non-network (13C)	400	60	70	20	0	80
Pair 14 (NY)	Network (14N)	400	50	80	20	0	100
	Non-network (14C)	500	50	80	10	0	70
Pair 15 (NY)	Network (15N)	300	50	40	60	0	70
	Non-network (9C)	400	40	40	50	0	70

Note. CA is California; FRPL is free or reduced-price lunch, and NY is New York City. School demographics from the 2010–11 Common Core of Data (CCD). To ensure school confidentiality, enrollment is rounded to the nearest 100 students, and percentages are rounded to the nearest 10%.

Details on Specific School Pairs:

- Schools 4N and 5N are in the same district, and we were able to recruit only a single non-network school in this district. The students in this non-network school were matched to students in both School 4N and School 5N.
- Schools 7N1 and 7N2 were associated with the same deeper learning network and resided on the same campus. Because the schools were small in size, we combined the students attending them and treated them as a single network school in the analyses, comparing it with 7C.
- Because of small sample sizes, Schools 10C1 and 10C2 (non-network schools) were combined and treated as a single non-network school. Both non-network schools served populations that were similar to Schools 10N and 11N (network schools), which were associated with the same deeper learning network. The propensity scores for Pairs 10 and 11 were based on a combined sample that included both Schools 10N and 11N (network schools) and Schools 10C1 and 10C2 (non-network schools) because of the limited sample size within the individual network and non-network schools. After the propensity scores had been computed, however, Pairs 10 and 11 were considered separate pairs for the purposes of the impact analysis.

[a] Because of missing data in the 2010–11 CCD, demographic information for this school come from the 2011–12 CCD, and free or reduced-price lunch information for this school came from 2011–12 enrollment data from the California Department of Education, 2011–12.

B. Student Samples

In each matched pair of schools, the study focused on four student cohorts. To account for preexisting differences between students attending network and non-network schools in our analyses, we restricted the sample to students who had data on Grade 8 characteristics, including middle school state standardized test scores, in the available district extant data (see Section III.C). This requirement restricted our sample to students who attended a district school in Grade 8, so our results may not generalize to students who attended a school in our sample in Grade 9 but attended a nondistrict middle school. The progression of these four cohorts through high school and after high school and the timing of key project activities are in Exhibit 2.3.

Exhibit 2.3. Study Timeline and Expected Academic Progression of Participating Grade 9 Cohorts

Study timeline	Academic year	Cohort 1	Cohort 2	Cohort 3	Cohort 4
Original study (Reports 1–3)	2007–08	9			
	2008–09	10	9		
	2009–10	11	10	9	
	2010–11	12	11	10	9
	2011–12	AHS1	12	11	10
	2012–13[a]	AHS2	AHS1	12	11
	2013–14	AHS3	AHS2	AHS1	12
Updated graduation and college enrollment outcomes (Reports 4–5)	2014–15	AHS4	AHS3	AHS2	AHS1
	2015–16	AHS5	AHS4	AHS3	AHS2
Follow-up study (Reports 6–7)	2016–17	AHS6	AHS5	AHS4	AHS3
	2017–18	AHS7	AHS6	AHS5	AHS4
	2018–19	AHS8	AHS7	AHS6	AHS5
	2019–20[b]	AHS9	AHS8	AHS7	AHS6

Note. AHSX refers to the number of years after expected high school graduation.

[a] Participants consented to participate in the study and administration of the high school survey.

[b] Last academic year of college enrollment and degree completion data that was included in NSC data requested in spring 2021.

Overall, AIR collected college enrollment and degree completion data for 17,075 study participants across the four Grade 9 cohorts. Because students from Cohort 1 and Cohort 2 had already graduated from high school by the time of our original primary data collection in spring 2013, only students from Cohort 3 and 4 (who were in Grade 11 or 12 at the time) were consented to participate in primary data collection. Analyses in Report 7 that included high school survey measures were therefore based on students in Cohort 3 and Cohort 4.

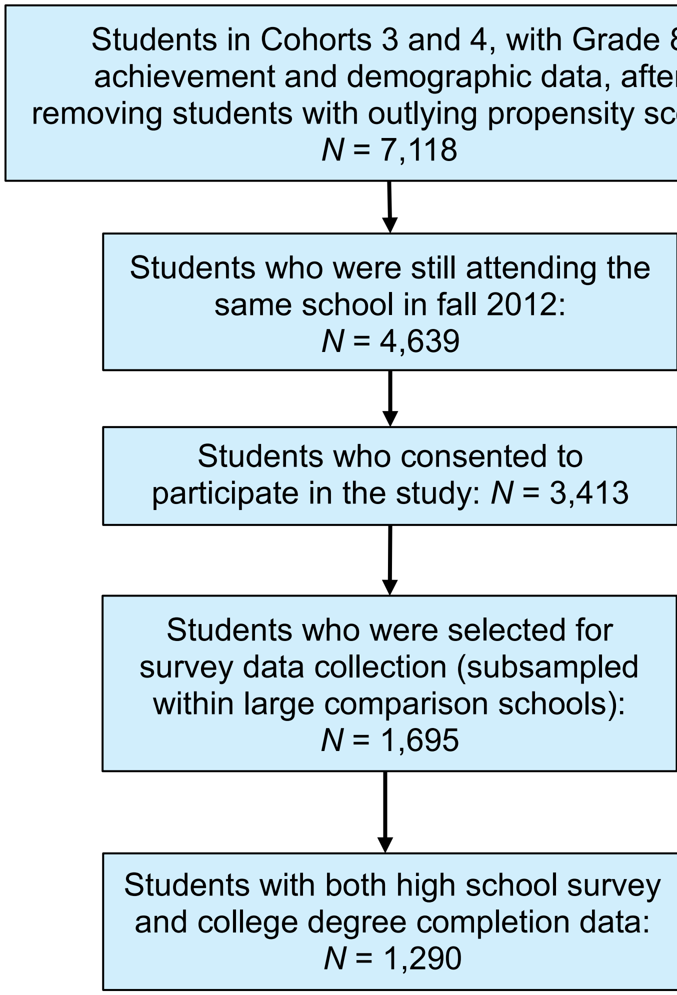
Sampling for Survey Data Collection

For primary data collection, our goal was to collect data from 260 students within each school pair (65 Grade 11 students and 65 Grade 12 students in the network and non-network schools). We selected student samples for primary data collection based on propensity score quintiles to ensure that we were sampling similar groups of students in each pair of schools. (For more information regarding the calculation of propensity scores, see Section IV.A.) The propensity score quintiles were defined based on the distribution of network students' estimated propensity scores—the conditional probability of being assigned to the treatment condition (network school enrollment) given a set of observable covariates (Rosenbaum & Rubin, 1983). To ensure that the students we sampled in matched non-network and network schools had similar background characteristics, we excluded non-network school students whose estimated propensity scores fell outside the region of "common support," which is loosely defined as the range of propensity scores of students who enrolled in the matched network school. In other words, we excluded students in non-network schools from the top propensity score stratum if they had unusually high propensity scores and from the lowest stratum if they had unusually low propensity scores.

Within each school pair, we sampled all consented students from network schools. However, because non-network schools in California tended to be larger in size, we subsampled consented students from these schools by randomly selecting students based on their propensity score quintile and the number of network students in the quintile. As a result, selected samples of network and non-network students had similar distributions of propensity scores within each matched pair of schools. Because the propensity scores reflect student background characteristics, the selected samples of network and non-network students also had similar characteristics.

Exhibit 2.4 illustrates the sample selection process for the school pairs in Report 7 with both high school survey and NSC data. The study began with 7,118 study participants who (a) entered Grade 9 in 2009–10 or 2010–11 (i.e., Cohort 3 or Cohort 4), (b) had nonmissing Grade 8 demographic and achievement data, and (c) had propensity scores that fell within the region of common support. Approximately 65% of these students (4,639 students) were still enrolled in the same high school in fall 2012, when the study team collected consent forms, and 3,413 participants consented to participate in the study. Sampling procedures resulted in a sample of 1,695 study participants who were selected to take a high school survey. Finally, as described in Section III.B, a high school survey administered in spring 2013 achieved a response rate of 76% (1,290 respondents) among students who attended schools for which we were able to obtain NSC data. The response rate was 79% among students who attended network schools and 75% among students who attended non-network schools.

Exhibit 2.4. Number of Students From the Initial Grade 9 Sample to the Data Collection Sample (Cohorts 3 and 4) Among Schools From Which We Obtained NSC Data



III. Data Sources and Measures

To address the research questions for this follow-up study, we collected college enrollment and degree completion data using the StudentTracker service at the NSC. We also used high school survey data collected in the original impact study. An overview of the data sources, including coverage across schools and students, is in Exhibit 3.1. Student-level administrative records from the participating districts also were collected for all students who entered Grade 9 in all four study cohorts to estimate propensity scores and include covariates in outcome models.

Exhibit 3.1. Outcome Data Sources and Sample Sizes for the Follow-Up Study

Data source	Description	Sample	Number of schools	Response rate
National Student Clearinghouse	Administrative records of college enrollment and degree completion	Students in Cohorts 1–4 in a subset of schools in the original study sample	24 schools; 12 school pairs	NA
High school survey	Measures students' self-reported opportunities to engage in deeper learning, as well as interpersonal and intrapersonal outcomes	Students in Cohorts 3 and 4 with parental consent and who were subsampled for data collection	22 schools; 11 school pairs	76% overall, 80% network students, and 73% non-network students

A. National Student Clearinghouse

AIR collected information on students' college enrollment and degree completion outcomes using the StudentTracker service at the NSC. The NSC is a nonprofit organization that collects student-level enrollment and degree completion information from postsecondary institutions in the United States. More than 3,600 institutions submit data to the NSC, accounting for approximately 98% of all students in public and private colleges across the United States.[3] Working closely with the districts participating in the study, we requested postsecondary data for students in Cohorts 1–4 within five districts, yielding 12 pairs of schools. We requested data for all students in the four cohorts who entered Grade 9 within our selected schools, including those who were not observed to have graduated from high school within the district.

For the follow-up study, we collected NSC data in spring 2021 that allowed us to measure the following postsecondary outcomes (captured up to 6 years after expected high school graduation) for all study participants:

[3] See https://www.studentclearinghouse.org/educational-organizations/studenttracker-for-educational-organizations/.

- Enrollment in any college, enrollment in a 2-year college, and enrollment in a 4-year college within the first year after expected high school graduation (three measures)
- Enrollment in any college, enrollment in a 2-year college, and enrollment in a 4-year college within the second year after expected high school graduation (three measures)
- Enrollment in college in both the fall of the first year and the fall of the second year after expected high school graduation (one measure)
- Regardless of the timing of the initial college enrollment, enrollment in two consecutive fall semesters or in two consecutive spring semesters within 6 years after expected high school graduation (one measure)
- Regardless of the timing of the initial college enrollment, enrollment in college for two, three, or four consecutive semesters (not counting summers) within 6 years after expected high school graduation (three measures)
- Completion of an associate's degree or certificate within 2, 3, or 4 years after expected high school graduation (three measures)
- Completion of a bachelor's degree within 4, 5, or 6 years after expected high school graduation (three measures)

B. High School Survey

As part of the survey development process, the high school survey was piloted in six network schools in spring 2012. To test the reliability of survey constructs and the survey administration processes, we subsampled 30 consented students from each high school grade to take the student survey. Items were added, dropped, or reworded based on findings from the pilot.

For the research study, high school surveys were administered in spring 2013, when respondents were expected to be in Grades 11 and 12. At most schools, the research team administered the surveys.[4] All schools were given the option of administering an online survey; paper surveys were administered in 18 schools, and students took online surveys in four schools. The student survey included items that measured opportunities to experience instruction focused on different dimensions of deeper learning and the competencies expected to result from exposure to deeper learning. For the follow-up study, we focused on nine measures of opportunities for deeper learning and eight measures of students' interpersonal and intrapersonal competencies that were measured in the high school survey. Survey items

[4] AIR staff were not present for survey administration in one school because of scheduling issues. In addition, students in two schools who were not present for the first survey administration were asked to complete the online survey on their own time. AIR staff were not present for these makeup sessions.

associated with these high school survey measures follow. A complete version of the high school survey can be found on the study's website.

Each survey item had four response options. For example, the items that measured opportunities for deeper learning had the following response options: None of my classes within the academic year (coded 0); one of my classes within the academic year (coded 1); two of my classes within the academic year (coded 2); and three or more of my classes within the academic year (coded 3). We estimated construct scores from the item-level responses with an ordered logit Rasch model (Yen, 1986), implemented with the WINSTEPS software package. The resulting Rasch scale scores are in the logit metric and have both negative and positive values. The value of zero is anchored to the average difficulty of the items included in the scale. In general, a student with a positive score tended to respond favorably (i.e., choosing the highest or second highest response option) on average. A student with a negative score tended to respond negatively (i.e., choosing the lowest or second lowest response option) on average. The sample on which we calculated Rasch scores for each scale was restricted to students with missing data for no more than half of the items within the scale. Less than 5% of the students within each school had missing data on each scale, except for one non-network school, in which a technological glitch during survey administration caused all items from the first half of the survey to be deleted.[5] High school survey measures used in Report 7 were standardized among the full sample of survey respondents, so the resulting scales had a mean of 0 and a standard deviation of 1.

Opportunities for Deeper Learning

Opportunities for Assessments Aligned With Deeper Learning

(Source: All original items)

Rasch reliability = .77; Cronbach's alpha = .86

Still thinking about the teachers of your English, math, science, and social studies classes this year, for how many of these classes is each statement true? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

My teacher gives tests about facts that we studied in class.
My teacher gives tests at the beginning of a unit to see how much we already know.
My teacher gives tests that require us to use different sources of information for our answers.
My teacher gives us points on a test or homework for how we solved a problem, not just whether we got the right answer.
My teacher asks us to put together a portfolio of many different examples of our school work.

[5] In one of the four schools in which the survey was administered online, a computer glitch deleted students' responses to the first half of the survey as soon as they advanced to the second half of the survey. We corrected the computer issue and asked students to retake the student survey, but only a small number of students retook the survey.

My teacher evaluates us on how well we work in groups.
My teacher asks us to evaluate ourselves on our class work.
My teacher asks us to explain our thinking.
My teacher has conferences with just me (not with my parents) so I can talk about what I'm learning in class and how well I'm doing.

Opportunities for Complex Problem Solving

(Source: Adapted from the National Survey of Student Engagement [NSSE], 2011)

Rasch reliability: .90; Cronbach's alpha: .93

In how many of your English, math, science, and social studies classes this year do you do the following? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

I analyze an idea, experience, theory, or story by examining its various parts.
I combine many ideas and pieces of information into something new and more complex.
I judge the value and reliability of an idea.
I use ideas or concepts from one class to help solve a problem in another classroom.

Opportunities for Complex Problem Solving in English Language Arts

(Source: Consortium on Chicago School Research [CCSR], 2007)

Rasch reliability: .83; Cronbach's alpha: .89

Think about your English classes you've taken this year. In these classes, how often do you do the following? [Never, Some of the time, Most of the time, All of the time]

I discuss my point of view about something I've read.
I discuss connections between what we are reading in class and real-life people or situations.
I discuss how culture, time, or place affects an author's writing.
I explain how writers use tools like symbolism and metaphor to communicate meaning.
I improve a piece of writing as a class or with partners.
I debate the meaning of what we are reading in class.

Opportunities for Complex Problem Solving in Mathematics

(Source: CCSR, 2007)

Rasch reliability: .71; Cronbach's alpha: .76

Now just think about your math classes this year. In these classes, how often do you do the following? [Never, Some of the time, Most of the time, All of the time]

I write a few sentences to explain how I solved a math problem.
I write a math problem for other students to solve.
I discuss possible solutions to problems with other students.
I use math to solve real-world problems.
I solve a problem with multiple steps that take more than 20 minutes.

Opportunities for Complex Problem Solving in Science

(Source: Original)

Rasch reliability: .86; Cronbach's alpha: .91

Now just think about your science classes you've taken this year. In these classes, how often do you do the following? [Never, Some of the time, Most of the time, All of the time]

I form hypotheses by asking questions and defining problems.
I create physical models representing scientific ideas.
I plan and carry out experiments.
I interpret data and explain what the results mean.
I use equations to help me analyze data or solve a problem.
I use data to support a hypothesis or argument.
I am required to judge the value and quality of information.

Opportunities for Creative Thinking

(Source: Created by the study team)

Rasch reliability: .79; Cronbach's alpha: .88

Still think about your English, math, science, and social studies classes this school year. For how many of these classes is each statement true? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

I am encouraged to come up with new and different ideas.
I need to think of original solutions to problems.
I am asked to come up with new ways to do things.
I am challenged to create new ideas.
I have to use my imagination.

Opportunities to Communicate

(Source: Created by the study team, based on the Common Core State Standards)

Rasch reliability: .83; Cronbach's alpha: .90

How many of your teachers (in your core academic subjects) this year ask you to do the following? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

I write for different purposes (for example, to explain or to persuade).
I write for different audiences.
I write and revise a piece of writing through multiple drafts.
I use technology and the Internet to write and get feedback on our writing (for example, on a message board or blog).
I write what I want in a journal, diary, or blog at least once a week.
I lead a group or class discussion.
I share my opinions in a class discussion.
I give presentations with visual aids, such as pictures, videos, charts, or graphs.
I give presentations.
I give presentations for different types of people, such as other students, parents, or people outside of school.
I discuss how well other students present their ideas in presentations.
I use information from different types of sources, such as videos, pictures, graphs, charts, and presentations.

Opportunities to Collaborate

(Source: Various, listed beside each question)

Rasch reliability: .69; Cronbach's alpha: .93

In how many of your core academic classes this year do you do each of the following? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

I work with other students on projects during class. *(NSSE, 2011)*
I work on assignments with my classmates outside of class. *(NSSE, 2011)*
I work in groups of two to six students. *(New York City, 2011)*
I need to work with others to do well in class. *(Original)*
Students review and discuss each other's work. *(Akey, 2006)*
Students help each other learn. *(Akey, 2006)*
Students ask questions and give feedback when others present their work in class. *(Akey, 2006)*

Students review what they've learned with one another. *(Akey, 2006)*
Students speak about their work in front of the class. *(Akey, 2006)*

Opportunities for Interdisciplinary Learning

(Source: Listed beside each question)

Rasch reliability: .78; Cronbach's alpha: .82

Still thinking about your English, math, science, and social studies classes this year, how often do you do the following? [Never, Some of the time, Most of the time, All of the time]

I work on a project that combines more than one subject (for example, science and literature). *(Original)*
I put together ideas or concepts from different subjects for assignments or discussions. *(NSSE, 2011)*
I attend a class that two teachers from different subjects teach together (for example, a math teacher and a science teacher). *(Buck Institute for Education [BIE], 2007)*
I use ideas or concepts from one class to help solve a problem in another class. *(Original)*

Opportunities to Learn How to Learn

(Source: Original, Measures of Effective Teaching[6] [MET] project)

Rasch Reliability: .52; Cronbach's Alpha: .78

Think about the teachers of your English, math, science, and social studies classes this year. For how many of these classes is each statement true? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

My teacher gives us activities to do, other than just listening to him or her. *(MET project)*
My teacher lets me test or try out my ideas to see if they work. *(MET project)*
My teacher helps me learn to use different sources of information.
My teacher asks me to think about how I learn best.

Opportunities for Real-World Connections

(Source: Various, listed beside each question)

Rasch reliability: .84; Cronbach's alpha: .89

[6] See http://www.metproject.org/.

Regarding the work you do for your core academic classes (such as English, math, science, and social studies) this year, in how many classes does the following happen? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

I make observations or collect data outside of the classroom for assignments. *(BIE, 2007)*
I interview or get information from family or community members. *(BIE, 2007)*
We connect what we are learning to life outside the classroom. *(CCSR, 2007)*
I work on helping solve real-world problems. *(CCSR, 2007)*
I find information for a project from sources outside of school. *(Original)*
We discuss how someone could use something we learned in school in a real job. *(Pace & Kuh, 1998)*
I can apply what I learn in class to my life outside of school. *(Walker & Fraser, 2005)*
I am able to pursue topics that interest me. *(Walker & Fraser, 2005)*
I work with real-world examples in class work. *(Original)*

Opportunities to Receive Feedback

(Source: Created by the study team, unless otherwise noted)

Rasch Reliability: .75; Cronbach's Alpha: .84

Think about your core academic classes this year and the feedback you receive about your work in those classes. For how many classes is each statement below true? [None of my classes, One of my classes, Two of my classes, Three or more of my classes]

My teacher gives me feedback on most of my work.
My teacher gives me specific suggestions about how I can improve my work. *(CCSR)*
I learn a lot from my teacher's feedback on my work.
I get useful feedback on my school work from other students.
I sometimes receive feedback on my work from someone other than the teacher or other students, such as my parents.
My teacher often asks me to revise my work after I get feedback.

Deeper Learning Competencies

Academic Engagement

(Source: Listed beside each question)

Rasch reliability = .74; Cronbach's alpha = .77

Regarding your core academic classes (English, math, science, and social studies) this year, to what extent do you agree with the following statements? [Strongly disagree, Disagree, Agree, Strongly agree]

The topics we are studying are interesting and challenging. (*CCSR, 2007)*
I am usually bored by classes or activities. (*CCSR, 2007)*
I usually look forward to classes or activities. (*CCSR, 2007)*
Sometimes I get so interested in my work I don't want to stop. (*CCSR, 2007)*
I often count the minutes until class ends. (*CCSR, 2007)*
I always prepare for class. (*Tinio, 2009*)
I ask questions when I don't understand the lesson. (*Tinio, 2009*)
I actively participate in group activities. (*Tinio, 2009*)
I am usually distracted by my classmates. (*Tinio, 2009*)
I cut class when I'm bored. (*Tinio, 2009*)

Collaboration Skills

(Source: Listed beside each question)

Rasch reliability = .83; Cronbach's alpha = .91

Now think about the group work you do for your classes. How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true] When I work with a group . . .

I tell the other members of my group when I think they are doing a good job. (*Huang et al., 2010)*
I make sure to be prepared and bring needed materials. (*Original)*
I remember to do my part of a group project without being reminded. (*Original)*
I finish my part of a group project on time. (*Original)*
I help keep my group focused. (*Original)*
I share my ideas with the group. (*Original)*
I help my group figure out and fix any problems we face. (*Original)*
I pay attention when my teammates talk. (*Original)*
I consider everyone's ideas. (*Original)*
I learn from other people in my group. (*Original)*

Creative Thinking

(Source: Original)

Rasch reliability: .77; Cronbach's alpha: .84

How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

I am able to come up with new and different ideas.
I like to think of original solutions to problems.
I come up with new ways to do things.
I am an original thinker.
I have a better imagination than my friends.

Perseverance

(Source: Duckworth & Quinn, 2009)

Rasch reliability = .79; Cronbach's alpha = .88

How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

I overcome setbacks to achieve important goals.
I am a hard worker.
I finish what I begin.
I achieve goals even if they take a long time.
I do a careful and thorough job. *(Original)*

Locus of Control

(Source: Levenson, 1981)

Rasch reliability = .73; Cronbach's alpha: .83

How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

I believe that whether or not I get to be a leader depends mostly on my ability.
When I make plans, I am almost certain to make them work.
I believe that I can pretty much determine what will happen in my life.
I believe that when I get what I want, it's usually because I worked hard for it.
I believe that my life is determined by my own actions.

Motivation to Learn

(Source: Pintrich & de Groot, 1990)

Rasch reliability: .75; Cronbach's alpha: .81

Think about the work you are doing in your classes this year. How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

It is important for me to learn what is being taught in my classes.
I think that what I am learning in my classes is useful for me to know.
I think what I am learning in my classes is interesting.
I prefer class work that is challenging so I can learn new things.
I try to learn from my mistakes in my schoolwork.

Self-Management

(Source: Listed beside each question)

Rasch reliability = .81; Cronbach's alpha = .85

How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

I set goals for doing better in school. *(Culture of Excellence & Ethics Assessment, 2019)*
I make a to-do list every day. *(Xue & Sun, 2011)*
I make schedules to help myself finish tasks on time. *(Xue & Sun, 2011)*
I finish my tasks on time. *(Xue & Sun, 2011)*
I get all the help I can to help me reach my goals. *(Xue & Sun, 2011)*
I set long-term goals for myself. *(Xue & Sun, 2011)*
I can find the information I need to learn on my own. *(Pace & Kuh, 1998)*
I feel good about my ability to learn whatever I want or need to know. *(Learning Point Associates, 2012)*
I can learn effectively on my own. *(NSSE, 2011)*
I feel like I am in charge of what I learn. *(Learning Point Associates, 2012)*

Self-Efficacy

(Source: Chen et al., 2001)

Rasch reliability = .84; Cronbach's alpha = .91

How often are the following statements true about you? [Never or almost never true, Sometimes true, Usually true, Always or almost always true]

I believe I will be able to reach my goals.
I know I can complete difficult tasks.
I believe I can do whatever I decide to do.
I believe I will be able to overcome challenges.
I know I can do many different things well.
Compared to most other people, I can do most tasks very well.
Even when things are tough, I can perform quite well.

C. Student Background Data (Extant Data)

We obtained student-level administrative records from the participating districts. These records contained data on student characteristics measured in Grade 8 and Grade 9. We used the record data to identify students to be included in our samples (i.e., first-time Grade 9 students) and to incorporate covariates in our analyses. Our study schools were in multiple school districts, so consistent data were not available for all study schools. However, because school pairs were constructed within a district, we had the same set of student background characteristics for the two schools in any given pair.[7] Exhibit 3.2 lists the student background data we received from districts and details how many school pairs had each data element. As the exhibit indicates, we had two measures of student socioeconomic background: parents' education and students' free or reduced-price lunch status. In impact models, we used a single measure of low socioeconomic status. Among schools with information about eligibility for free or reduced-price lunch, students who were eligible for free or reduced-price lunch were identified as low socioeconomic status. Among schools with information on parental education, students with parents with a high school education or less were identified as low socioeconomic status.

[7] One pair of schools contained a network and a non-network school in neighboring districts. The data elements available across the two districts were very similar.

Exhibit 3.2. Description of Student Background Data From Extant District Data

Measure	Description	Number of school pairs with available data
Female	Dichotomous indicator of students' gender	15
Race/ethnicity	Dichotomous indicators created for African American, Hispanic, White, Asian, and "other" races	15
Parents' education	Categorical measure of parental education—specifically, the highest level of education obtained by either parent—using the following categories: some high school, high school diploma, some college, college degree, higher degree (above BA), and declined to report parents' education (varies slightly by district)	6
Free or reduced-price lunch status	Dichotomous indicator of whether student was eligible for the free or reduced-price lunch program, typically in Grade 8	9
English learners	Dichotomous indicator of whether the student was identified as an English learner, typically in Grade 8	15
Individualized education program	Dichotomous indicator of whether the student had an individualized education program, typically in Grade 8	15
Prior achievement in English language arts	Standardized test score in English language arts prior to entering high school, from Grade 8	13
Prior achievement in mathematics	Standardized test score in mathematics prior to entering high school, from Grade 8, including indicators for math test subject where relevant; standardized using the state mean and standard deviation for each year and grade level	13

IV. Analytic Methods

In this section, we describe the calculation of analysis weights as well as the analysis methods employed to estimate (a) the impact of attending deeper learning network schools (impact analyses) and (b) the relationships between students' opportunities for deeper learning and deeper learning competencies in high school and college degree completion outcomes (correlational analyses).

A. Weighting

We applied weights to statistical analyses to reflect two features of the study's design. First, we applied propensity score weights to account for measured pre–high school characteristics (including both demographic characteristics and Grade 8 achievement test scores) related to the decision to enroll in a deeper learning high school and likely related to student outcomes. Second, we applied a "survey weight" that accounted for attrition during high school, sampling, and survey nonresponse. Survey weights were estimated as inverse probability weights so that results for the students from whom we collected high school survey data would be representative of the students who entered sampled network and non-network schools in Grade 9 from the two cohorts that participated in the high school survey (i.e., who entered Grade 9 in 2009–10 or 2010–11). In this section, we describe both the propensity score weight and the survey weight in greater detail.

Propensity Score Weights: Weights for Student Selection Into Network Schools

Students were not randomly assigned to attend network and non-network schools, so network and non-network school students may not have had equivalent characteristics when entering high school. These preexisting student differences mean that any claims about a network school's effects on student outcomes could be biased if based on direct comparisons between network and non-network school students. To account for these preexisting differences, we used inverse probability of treatment weighting (IPTW), a propensity-score-based method for selection bias adjustment (Hirano et al., 2003). Assuming the measured student background characteristics accurately capture the important preexisting differences between network and non-network students, IPTW allows us to obtain valid estimates about what network students would have experienced had they attended non-network schools.

Specifically, we used generalized boosted regression (McCaffrey et al., 2004), which iteratively tried various combinations of student background covariates to predict the probability of enrolling in a deeper learning network school in Grade 9. The algorithm searches for the combination of covariates that minimizes the differences in measured characteristics between students who enrolled in a network school and students who enrolled in a non-network school when the latter are weighted by the inverse probability of enrolling in a network school. We used the twang package in the R statistical program to execute the generalized boosted regression. Following the recommendations set forth by the package authors (Ridgeway et al., 2013), we set the interaction depth to 4, shrinkage to .0005, and bagging to .50. We used the following equation to estimate students' propensity score (p_i), which is their predicted probability of attending a network school instead of a non-network school in Grade 9, given the measured student characteristics (X_i):

$$\ln\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_{0jk} + \beta_{1jk}\boldsymbol{X_{ijk}}$$

where $\boldsymbol{X_{ijk}}$ represents the student characteristics listed in Exhibit 3.2, students' entering the Grade 9 cohort, and school pair fixed effects. The estimated propensity scores were then used to calculate "average treatment effect on the treated" (ATT) weights for study participants using the equation:

$$w_{ijk} = T_{ijk} + \left(1 - T_{ijk}\right)\frac{p_{ijk}}{1 - p_{ijk}}$$

where T_{ijk} equals 1 for students attending a network school and 0 for students attending a non-network school. Using this equation, the ATT weight had a value of 1 for all students attending a network school, and the non-network school students were weighted to represent the network school students (with study participants with larger propensity scores given larger weights) to facilitate estimation of the average treatment effect on the treated.

The ATT weight was applied to analyses that examined the impact of attending a deeper learning network school on college enrollment, persistence, and degree completion outcomes. To examine the performance of the ATT weight, we assessed the degree to which network and non-network school students had similar student background characteristics after applying the ATT weight and accounting for the nesting of students in school pairs. In the baseline equivalence results in Exhibit 4.1 (for the full sample) and Exhibit 4.2 (separately for students in California and New York City), we used the Cox index to estimate effect sizes for binary characteristics and Hedge's g to estimate effect sizes for continuous achievement test scores.

Exhibit 4.1 shows that, except for English learner (EL) status and identification as a student with an individualized education program (IEP), effect sizes of differences between students who attended network schools and students who attended non-network schools did not exceed 0.05 standard deviation, which is the low threshold for baseline imbalance (What Works Clearinghouse, 2020). Effect sizes for EL status and IEP status fall above the 0.05 standard deviation threshold but below the 0.25 standard deviation threshold. As such, these covariates are included in impact models. We also generally observed baseline equivalence when we separately examined students who attended high school in California and students who attended high school in New York City (see Exhibit 4.2), with student characteristics demonstrating greater similarity in California. Similar to our observation in the full sample, among students who attended high school in New York City, effect sizes for EL status and IEP status fall above the 0.05 standard deviation threshold but below the 0.25 standard deviation threshold.

Exhibit 4.1. Network and Non-Network Student Characteristics for the Impact Analysis Sample

Student characteristics	Network mean	Non-network mean	SMD
Average standardized Grade 8 English language arts test score[a]	-0.05	-0.03	-0.02
Average standardized Grade 8 mathematics test score[a]	-0.04	-0.03	-0.01
Percentage in the 2007–08 cohort	24.6%	24.8%	-0.01
Percentage in the 2008–09 cohort	26.7%	27.0%	-0.01
Percentage in the 2009–10 cohort	24.6%	24.7%	0.00
Percentage in the 2010–11 cohort	24.2%	23.9%	0.01
Percentage female	56.5%	55.9%	0.01
Percentage Black	32.1%	32.3%	0.00
Percentage Hispanic	47.9%	47.3%	0.01
Percentage White	14.6%	14.6%	0.00
Percentage Asian/other race	5.4%	5.7%	-0.04
Percentage low socioeconomic status[b]	62.8%	63.1%	-0.01
Percentage students with an individualized education program	6.1%	5.3%	0.09
Percentage English learner students	24.2%	21.4%	0.09

Note. SMD = standardized mean difference. Adjusted group averages for students who attended non-network schools were calculated using a weighted two-level regression model with students nested in school pairs. Test scores were standardized within the original cohort sample by state and Grade 9 cohort. SMDs were calculated using the Cox Index for binary characteristics and Hedge's *g* for continuous achievement test scores. Results are based on the sample of 17,075 study participants with outcome data (3,547 students who attended network schools and 13,528 students who attended non-network schools).

[a] Excludes two school pairs without prior achievement data. [b] Includes students who qualified for free or reduced-price lunch or whose parents had a high school education.

Exhibit 4.2. Network and Non-Network Student Characteristics for the Impact Analysis Sample, Separately for Students in California and Students in New York City

Student characteristics	California			New York City		
	Network mean	Non-network mean	SMD	Network mean	Non-network mean	SMD
Average standardized Grade 8 English language arts test score[a]	-0.12	-0.12	0.00	0.00	0.03	-0.03
Average standardized Grade 8 mathematics test score[a]	-0.14	-0.13	0.00	0.02	0.04	-0.02
Percentage in the 2007–08 cohort	21.0%	21.9%	-0.03	26.6%	25.7%	0.03
Percentage in the 2008–09 cohort	25.4%	25.2%	0.01	27.4%	27.1%	0.01
Percentage in the 2009–10 cohort	25.6%	25.2%	0.01	24.0%	25.0%	-0.03
Percentage in the 2010–11 cohort	28.1%	27.8%	0.01	22.0%	22.5%	-0.02
Percentage female	54.2%	54.3%	0.00	57.8%	56.7%	0.03
Percentage Black	14.9%	14.4%	0.02	41.8%	42.3%	-0.01
Percentage Hispanic	57.8%	57.6%	0.01	42.4%	41.6%	0.02
Percentage White	18.4%	18.6%	-0.01	12.5%	12.4%	0.01
Percentage Asian/other race	9.0%	9.4%	-0.03	3.3%	3.6%	-0.05
Percentage low socioeconomic status[b]	51.3%	52.4%	-0.03	69.3%	69.0%	0.01
Percentage students with an individualized education program	8.1%	8.2%	-0.01	5.0%	3.8%	0.18
Percentage English learners	26.1%	25.3%	0.02	23.1%	17.8%	0.20

Note. SMD = standardized mean difference. Adjusted group averages for students who attended non-network schools were calculated using a weighted two-level regression model with students nested in school pairs. Test scores were standardized within the original cohort sample by state and Grade 9 cohort. SMDs were calculated using the Cox Index for binary characteristics and Hedge's *g* for continuous achievement test scores. Results for students who attended high school in California are based on the sample of 11,475 study participants with outcome data (1,272 who attended network schools and 10,203 who attended non-network schools). Results for students who attended high school in New York City are based on the sample of 5,600 study participants with outcome data (2,275 who attended network schools and 3,325 who attended non-network schools).
[a] Excludes two school pairs without prior achievement data. [b] Includes students who qualified for free or reduced-price lunch or whose parents had a high school education.

Survey Weights: Weights for Student Persistence, Consent, and Survey Nonresponse

For the original *Study of Deeper Learning*, we selected a subsample of study participants to take the high school survey. To be selected for active data collection, study participants must have (a) persisted in the same high school until fall 2012 and (b) consented to participate in active data collection during the 2012–13 school year. Both attrition (i.e., leaving the high school prior to fall 2012) and nonconsent had the potential to bias the study sample because the characteristics of students who consented to participate in the study may not resemble the characteristics all students who entered these Grade 9 cohorts. In addition, to limit the number of survey respondents from large non-network high schools, no more than 260 consented students from each school pair were sampled for survey data collection.[8] Finally, nonresponse to the high school survey had the potential to introduce bias into the analytic sample because study participants were excluded from analyses if they had missing data on relevant survey measures. Approximately 76% of sampled students responded to the high school survey.

Aligning with prior research on survey response bias, descriptive statistics indicated that the subset of students who responded to the high school survey differed in measured characteristics from the full sample of students in those entering Grade 9 cohorts. To account for sample attrition, nonconsent, sampling, and nonresponse, we estimated survey weights using students' Grade 8 demographic characteristics and achievement test scores. This survey weight was applied to analyses estimating the relationships between opportunities for deeper learning and deeper learning competencies during high school and college degree completion outcomes.

To calculate survey weights, we used the *twang* package in the *R* statistical program to estimate generalized boosted regression models, as we described for the propensity score weights previously. These models estimated each student's probability of providing survey data. Survey weights were calculated as the inverse of the estimated probabilities of responding to the high school survey. In other words, study participants with lower probabilities of responding to the survey were given greater weight in analyses. With these weights, students in the analytic sample for the correlational analyses were weighted to represent the cohorts of entering Grade 9 students in sampled schools.

In Exhibit 4.3, we present descriptive statistics for (a) the sample of students who entered Grade 9 in sampled schools in 2009–10 and 2010–11 and for whom we were able to collect NSC data and (b) the subsample of these participants who responded to the high school survey. We provide descriptive information about the subsample of high school survey respondents both before and after applying the survey weights. In general, the application of survey weights

[8] Because network schools were smaller in size than non-network schools, we administered the survey to all consented network students. In school pairs in which network schools had fewer than 130 consented students within Grade 11 or Grade 12, we oversampled consented students in the matched non-network schools to achieve the target sample size of 260 students within each matched school pair. In large non-network schools with large numbers of consented students, we sampled consented students based on their propensity score strata (quintiles defined by the distribution of the matched network school).

reduced observed differences between the subsample of participants who responded to the high school survey and the original sample of Grade 9 students.

Exhibit 4.3. Student Characteristics in Cohort and Analytic Samples for Correlational Analyses, Before and After Applying Survey Weights

Student characteristic	Cohort sample mean (*n* = 9,118)	Respondents to the high school survey (*n* = 1,290)	
		Unweighted mean	Weighted mean
Average standardized Grade 8 English language arts test score[a]	0.032	0.183	0.035
Average standardized Grade 8 mathematics test score[a]	0.033	0.141	0.011
Percentage member of the younger cohort	49.8%	51.9%	48.4%
Percentage female	50.5%	53.3%	51.2%
Percentage Black	12.7%	13.1%	12.4%
Percentage Hispanic	46.5%	52.6%	50.8%
Percentage White	28.1%	25.1%	25.7%
Percentage Asian/other race	12.7%	9.1%	11.0%
Percentage low socioeconomic status[b]	49.6%	59.9%	51.3%
Percentage students with an individualized education program	7.9%	5.7%	7.5%
Percentage English learner students	25.0%	29.5%	25.1%

Note. Test scores were standardized within the original cohort sample by state and Grade 9 cohort.
[a] Excludes two school pairs without prior achievement data. [b] Includes students who qualified for free or reduced-price lunch or whose parents had a high school education.

B. Statistical Models

Impact Models

To estimate the effects of enrolling in a deeper learning network school instead of a non-network school, we estimated hierarchical linear models with study participants nested within school pairs. The analysis method is considered doubly robust (Funk et al., 2011) because it accounts for observed differences in network and non-network students through (a) propensity score weighting (i.e., the IPTW) and (b) regression-based covariate adjustment. If either of the two adjustment methods accurately accounts for differences in student baseline characteristics, then we can obtain valid estimates of the impact of attending network schools on student outcomes. However, because the network schools in this study were purposefully selected to be moderately or high implementing (according to their networks) and to meet other criteria, results from this study cannot be generalized to all schools within the participating networks.

Impact analyses used the following weighted two-level hierarchical linear model (presented in reduced form), with students nested within school pairs:

$$Y_{ij} = \gamma_{00} + \gamma_{01}T_{ij} + \gamma_{10}\boldsymbol{X_{ij}} + u_{0j} + r_{ij}$$

where Y_{ij} is a given participant outcome for student *i* in school pair *j*; T_{ij} is a dichotomous indicator for whether the student enrolled in the network school (T_{ij} = 1) or the non-network school (T_{ij} = 0) in the fall of Grade 9; and $\boldsymbol{X_{ij}}$ is a vector of available student background characteristics listed in Exhibit 3.2, as well as dichotomous indicators for incoming Grade 9 cohorts.

The main parameter of interest is γ_{01}, which is the estimated impact of attending a network school instead of the matched non-network school. Using hierarchical linear models for binary outcomes, estimates of γ_{01} can be interpreted as percentage point differences between the two study groups.[9]

Correlational Analysis Models

The impact model was modified to estimate the relationships between opportunities for deeper learning and deeper learning competencies in high school and college degree completion outcomes. We estimated relationships between high school measures and outcome measures using hierarchical linear models, with study participants nested within the schools they entered in Grade 9. These models applied a survey weight that accounted for nonresponse to the high school survey, and they controlled for student background characteristics. We estimated a separate model for each high school survey measure (M_{ij}), and high school survey measures were centered around the school mean. Correlational analyses used the following weighted two-level hierarchical linear model (presented in reduced form):

$$Y_{ij} = \gamma_{00} + \gamma_{10}M_{ij} + \gamma_{20}\boldsymbol{X}_{ij} + u_{0j} + r_{ij}$$

The main parameter of interest for these models is γ_{10}, which is the relationship between the specified high school survey measure and the degree completion outcome. To limit the number of statistical models and reduce the likelihood of a Type 1 error (i.e., incorrectly identifying a nonsignificant relationship as being statistically significant), we limited these analyses to two key outcomes that represent 150% of the normative time to degree completion: completion of

[9] To assist with the interpretation of findings, we estimated hierarchical linear models for binary outcomes. Alternative analyses used hierarchical generalized linear models with a logit link function for binary outcomes. The results of these analyses do not differ substantively from the analyses presented in Report 7.

an associate's degree or certificate within 3 years after expected high school graduation and completion of a bachelor's degree within 6 years after expected high school graduation (see detailed results in Section V).

Qualitative Analysis

The research team systematically analyzed data collected through telephone interviews with 20 study participants to identify themes related to respondents' post-high-school experiences and how their high schools prepared them, or could have better prepared them, for life after high school. When sufficient numbers of responses were available, we compared responses from participants who attended network schools and participants who attended non-network schools to identify possible differences.

Handling Missing Data

Because students must have had Grade 8 administrative data to be included in the original *Study of Deeper Learning*, few study participants have missing data on individual background characteristics. However, because two school pairs primarily served recent immigrants to the United States, and because these students were exempt from participating in state testing, prior achievement test scores were missing for participants within two school pairs. For Grade 8 achievement test scores in mathematics and English language arts, we imputed a value of 0 for the school pairs with missing data. Because analyses account for the clustering of participants within schools or school pairs, and a missing data indicator would be collinear with school pair membership, models did not include dummy variables for missing data.

Because the NSC collects data for more than 98% of the students enrolled in public and private colleges, we assumed that students for whom the NSC did not locate college enrollment records did not enroll in college. Therefore, the outcome measures included in Report 7 do not contain missing values. However, correlational analyses included only those participants who had nonmissing data on relevant high school survey measures of opportunities for deeper learning and deeper learning competencies. Sample sizes associated with each correlational analysis model are presented with results of correlational analyses in Exhibits 5.4 to 5.7 in Section V.

V. Detailed Results

In this section, we provide supplemental tables presenting detailed information for the results described in Report 7.

Exhibit 5.1. The Impact of Attending a Deeper Learning Network School on College Enrollment Outcomes, Overall and Separately for California and New York City

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
Overall Sample					
Enrolled in college by the end of the first year after expected high school graduation					
Any college	0.01	0.01	48%	46%	.313
2-year college	-0.03	0.03	24%	27%	.318
4-year college	0.05	0.03	25%	21%	.089
Enrolled in college by the end of the second year after expected high school graduation					
Any college	0.02	0.01	55%	53%	.122
2-year college	-0.02	0.04	32%	34%	.616
4-year college	0.05	0.03	27%	22%	.078
California Sample					
Enrolled in college by the end of the first year after expected high school graduation					
Any college	0.03	0.02	56%	54%	.193
2-year college	-0.01	0.03	36%	37%	.780
4-year college	0.04	0.03	23%	19%	.239
Enrolled in college by the end of the second year after expected high school graduation					
Any college	0.05	0.02	64%	60%	.026
2-year college	0.03	0.03	48%	45%	.360
4-year college	0.05	0.04	25%	20%	.222

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
New York City Sample					
Enrolled in college by the end of the first year after expected high school graduation					
Any college	0.00	0.02	43%	43%	.925
2-year college	-0.05	0.05	17%	22%	.319
4-year college	0.05	0.04	26%	21%	.205
Enrolled in college by the end of the second year after expected high school graduation					
Any college	0.00	0.02	49%	49%	.843
2-year college	-0.05	0.06	24%	29%	.390
4-year college	0.06	0.04	29%	23%	.188

Note. The percentages for network school students are unadjusted percentages; the percentages for non-network students are model-adjusted percentages. Overall results are based on the full sample of 17,075 study participants within 12 matched pairs of network and non-network schools (3,547 network school students and 13,528 non-network school students). Results for students who attended high school in California are based on the sample of 11,475 study participants with outcome data (1,272 network school students and 10,203 non-network school students). Results for students who attended high school in New York City are based on the sample of 5,600 study participants with outcome data (2,275 network school students and 3,325 non-network school students).

Exhibit 5.2. The Impact of Attending a Deeper Learning Network School on College Persistence Outcomes, Overall and Separately for California and New York City

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
Overall Sample					
Enrolled in college in both fall of the first year and fall of the second year after expected high school graduation	0.00	0.01	31%	31%	.978
Regardless of the timing of initial college enrollment, enrolled in two consecutive fall (or spring) semesters	0.01	0.01	41%	40%	.430
Regardless of the timing of initial college enrollment, enrolled in college for two consecutive semesters	0.02	0.02	51%	49%	.277
Regardless of the timing of initial college enrollment, enrolled in college for three consecutive semesters	0.01	0.01	41%	41%	.578
Regardless of the timing of initial college enrollment, enrolled in college for four consecutive semesters	0.01	0.02	36%	35%	.657
California Sample					
Enrolled in college in both fall of the first year and fall of the second year after expected high school graduation	-0.02	0.02	36%	38%	.370
Regardless of the timing of initial college enrollment, enrolled in two consecutive fall (or spring) semesters	0.01	0.01	51%	50%	.428
Regardless of the timing of initial college enrollment, enrolled in college for two consecutive semesters	0.03	0.02	60%	57%	.154
Regardless of the timing of initial college enrollment, enrolled in college for three consecutive semesters	0.00	0.02	48%	48%	.989
Regardless of the timing of initial college enrollment, enrolled in college for four consecutive semesters	-0.01	0.03	42%	43%	.707

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
New York City Sample					
Enrolled in college in both fall of the first year and fall of the second year after expected high school graduation	0.01	0.01	28%	27%	.633
Regardless of the timing of initial college enrollment, enrolled in two consecutive fall (or spring) semesters	0.00	0.02	36%	35%	.802
Regardless of the timing of initial college enrollment, enrolled in college for two consecutive semesters	0.00	0.02	46%	46%	.833
Regardless of the timing of initial college enrollment, enrolled in college for three consecutive semesters	0.01	0.02	38%	37%	.636
Regardless of the timing of initial college enrollment, enrolled in college for four consecutive semesters	0.02	0.03	32%	30%	.522

Note. All college persistence outcomes were measured within 6 years after expected high school graduation. The percentages for network school students are unadjusted percentages; the percentages for non-network students are model-adjusted percentages. Overall results are based on the full sample of 17,075 study participants within 12 matched pairs of network and non-network schools (3,547 network school students and 13,528 non-network school students). Results for students who attended high school in California are based on the sample of 11,475 study participants with outcome data (1,272 network school students and 10,203 non-network school students). Results for students who attended high school in New York City are based on the sample of 5,600 study participants with outcome data (2,275 network school students and 3,325 non-network school students).

Exhibit 5.3. The Impact of Attending a Deeper Learning Network School on College Degree Completion Outcomes, Overall and Separately for California and New York City

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
Overall Sample					
Completion of an associate's degree or certificate					
Within 2 years after expected high school graduation	0.00	0.00	1%	1%	.111
Within 3 years after expected high school graduation	-0.01	0.01	3%	4%	.203
Within 4 years after expected high school graduation	-0.01	0.01	5%	6%	.225
Completion of a bachelor's degree					
Within 4 years after expected high school graduation	0.00	0.01	8%	7%	.689
Within 5 years after expected high school graduation	0.01	0.01	13%	12%	.577
Within 6 years after expected high school graduation	0.01	0.01	16%	15%	.434
California Sample					
Completion of an associate's degree or certificate					
Within 2 years after expected high school graduation	0.00	0.00	1%	1%	.456
Within 3 years after expected high school graduation	-0.01	0.01	2%	3%	.240
Within 4 years after expected high school graduation	-0.02	0.01	3%	5%	.133
Completion of a bachelor's degree					
Within 4 years after expected high school graduation	-0.02	0.01	5%	7%	.201
Within 5 years after expected high school graduation	-0.03	0.01	10%	13%	.058
Within 6 years after expected high school graduation	-0.03	0.01	14%	17%	.045

Outcome	Coefficient	Standard error	Probability for network students	Probability for non-network students	*p* value
New York City Sample					
Completion of an associate’s degree or certificate					
Within 2 years after expected high school graduation	0.00	0.00	1%	1%	.424
Within 3 years after expected high school graduation	0.00	0.01	4%	4%	.798
Within 4 years after expected high school graduation	0.00	0.01	6%	6%	.999
Completion of a bachelor’s degree					
Within 4 years after expected high school graduation	0.02	0.01	9%	8%	.118
Within 5 years after expected high school graduation	0.02	0.01	14%	11%	.002
Within 6 years after expected high school graduation	0.03	0.01	17%	14%	.001

Note. The percentages for network school students are unadjusted percentages; the percentages for non-network students are model-adjusted percentages. Overall results are based on the full sample of 17,075 study participants within 12 matched pairs of network and non-network schools (3,547 network school students and 13,528 non-network school students). Results for students who attended high school in California are based on the sample of 11,475 study participants with outcome data (1,272 network school students and 10,203 non-network school students). Results for students who attended high school in New York City are based on the sample of 5,600 study participants with outcome data (2,275 network school students and 3,325 non-network school students).

Exhibit 5.4. Relationships Between Opportunities for Deeper Learning and Degree Completion Outcomes

Outcome	Coefficient	Standard error	*p* value	Sample size
Completion of an associate's degree or certificate within 3 years after expected high school graduation				
Assessments aligned with deeper learning	0.0%	0.008	.978	1,205
Opportunities for collaboration	-0.6%	0.008	.442	1,214
Opportunities for communication	-0.1%	0.008	.894	1,135
Opportunities for complex problem solving	0.9%	0.008	.246	1,139
Opportunities for creative thinking	0.8%	0.007	.260	1,205
Opportunities for feedback to students	-0.2%	0.007	.791	1,210
Opportunities for interdisciplinary learning	-0.2%	0.008	.753	1,199
Opportunities for learning how to learn	0.5%	0.007	.534	1,138
Opportunities for real-world connections	0.0%	0.006	.974	1,206
Completion of a bachelor's degree within 6 years after expected high school graduation				
Assessments aligned with deeper learning	0.8%	0.015	.597	1,205
Opportunities for collaboration	2.2%	0.020	.288	1,214
Opportunities for communication	0.4%	0.014	.796	1,135
Opportunities for complex problem solving	1.4%	0.010	.156	1,139
Opportunities for creative thinking	0.3%	0.013	.828	1,205
Opportunities for feedback to students	2.7%	0.013	.040	1,210
Opportunities for interdisciplinary learning	-0.9%	0.010	.370	1,199
Opportunities for learning how to learn	2.2%	0.010	.021	1,138
Opportunities for real-world connections	1.0%	0.014	.465	1,206

Note. Each coefficient represents the percentage point change in a given degree completion outcome associated with an increase of one standard deviation in a given deeper learning opportunity measure. Results are based on the subsample of study participants who entered Grade 9 in 2009–10 or 2010–11 within 19 sampled schools, and for whom we were able to collect both high school survey and National Student Clearinghouse data. Sample sizes vary across analyses due to missing data on relevant high school survey measures.

Exhibit 5.5. Relationships Between Deeper Learning Competency Outcomes and Degree Completion Outcomes

Outcome	Coefficient	Standard error	*p* value	Sample size
Completion of an associate's degree or certificate within 3 years after expected high school graduation				
Academic engagement	0.8%	0.007	.242	1,138
Collaboration skills	1.3%	0.007	.047	1,136
Creative thinking skills	0.6%	0.007	.435	1,133
Locus of control	1.5%	0.007	.038	1,203
Motivation to learn	1.3%	0.007	.059	1,136
Perseverance	1.3%	0.005	.004	1,134
Self-efficacy	0.8%	0.006	.180	1,203
Self-management	0.7%	0.007	.242	1,138
Completion of a bachelor's degree within 6 years after expected high school graduation				
Academic engagement	1.6%	0.010	.097	1,138
Collaboration skills	2.1%	0.011	.057	1,136
Creative thinking skills	1.3%	0.009	.175	1,133
Locus of control	3.8%	0.009	<.001	1,203
Motivation to learn	1.4%	0.014	.327	1,136
Perseverance	3.7%	0.008	<.001	1,134
Self-efficacy	3.3%	0.009	<.001	1,203
Self-management	3.6%	0.010	<.001	1,138

Note. Each coefficient represents the percentage point change in a given degree completion outcome associated with an increase of one standard deviation in a given deeper learning competency measure. Results are based on the subsample of study participants who entered Grade 9 in 2009–10 or 2010–11 within 19 sampled schools, and for whom we were able to collect both high school survey and National Student Clearinghouse data. Sample sizes vary across analyses due to missing data on relevant high school survey measures.

Exhibit 5.6. Relationships Between Opportunities for Deeper Learning and Degree Completion Outcomes—Separately for California and New York City

Outcome	California				New York City			
	Coefficient	Standard error	*p* value	Sample size	Coefficient	Standard error	*p* value	Sample size
Completion of an associate's degree or certificate within 3 years after expected high school graduation								
Assessments aligned with deeper learning	0.8%	0.008	.307	914	-3.9%	0.014	.006	291
Opportunities for collaboration	-0.3%	0.009	.707	921	-2.8%	0.018	.124	293
Opportunities for communication	0.0%	0.009	.972	843	-1.5%	0.017	.397	292
Opportunities for complex problem solving	1.0%	0.008	.203	846	-0.3%	0.027	.913	293
Opportunities for creative thinking	1.1%	0.008	.188	912	-1.2%	0.017	.464	293
Opportunities for feedback to students	0.4%	0.007	.549	917	-4.1%	0.024	.092	293
Opportunities for interdisciplinary learning	0.5%	0.007	.486	910	-5.1%	0.027	.055	289
Opportunities for learning how to learn	1.0%	0.006	.120	845	-3.1%	0.027	.244	293
Opportunities for real-world connections	0.2%	0.007	.790	913	-1.5%	0.010	.144	293
Completion of a bachelor's degree within 6 years after expected high school graduation								
Assessments aligned with deeper learning	1.0%	0.018	.590	914	-0.4%	0.029	.895	291
Opportunities for collaboration	2.0%	0.025	.422	921	2.0%	0.030	.516	293
Opportunities for communication	1.0%	0.017	.540	843	-2.7%	0.033	.421	292
Opportunities for complex problem solving	1.0%	0.011	.357	846	2.9%	0.024	.219	293
Opportunities for creative thinking	0.2%	0.014	.902	912	-0.8%	0.029	.773	293
Opportunities for feedback to students	3.1%	0.016	.049	917	-1.0%	0.022	.644	293
Opportunities for interdisciplinary learning	-0.1%	0.010	.890	910	-4.3%	0.025	.086	289
Opportunities for learning how to learn	2.3%	0.011	.033	845	-0.6%	0.024	.796	293
Opportunities for real-world connections	0.7%	0.016	.646	913	1.2%	0.021	.556	293

Note. Each coefficient represents the percentage point change in a given degree completion outcome associated with an increase of one standard deviation in a given deeper learning opportunity measure. Results are based on the subsample of study participants who entered Grade 9 in 2009–10 or 2010–11 within 19 sampled schools, and for whom we were able to collect both high school survey and National Student Clearinghouse data. Sample sizes vary across analyses due to missing data on relevant high school survey measures.

Exhibit 5.7. Relationships Between Deeper Learning Competency Outcomes and Degree Completion Outcomes—Separately for California and New York City

Outcome	California				New York City			
	Coefficient	Standard error	*p* value	Sample size	Coefficient	Standard error	*p* value	Sample size
Completion of an associate's degree or certificate within 3 years after expected high school graduation								
Academic engagement	0.8%	0.008	.324	846	0.5%	0.012	.683	292
Collaboration skills	1.3%	0.008	.093	844	1.2%	0.015	.448	292
Creative thinking skills	0.5%	0.008	.532	840	1.2%	0.026	.644	293
Locus of control	1.2%	0.007	.065	913	2.4%	0.028	.384	290
Motivation to learn	1.5%	0.007	.047	843	-0.2%	0.013	.901	293
Perseverance	1.3%	0.004	.004	841	0.8%	0.019	.658	293
Self-efficacy	0.7%	0.005	.205	913	0.7%	0.022	.735	290
Self-management	1.0%	0.008	.217	845	-2.2%	0.013	.092	293
Completion of a bachelor's degree within 6 years after expected high school graduation								
Academic engagement	1.7%	0.010	.096	846	0.8%	0.030	.805	292
Collaboration skills	1.7%	0.012	.169	844	5.2%	0.027	.052	292
Creative thinking skills	1.6%	0.011	.141	840	0.0%	0.039	.993	293
Locus of control	3.1%	0.008	<.001	913	6.8%	0.025	.006	290
Motivation to learn	1.2%	0.016	.457	843	3.5%	0.029	.226	293
Perseverance	3.6%	0.010	.001	841	5.0%	0.019	.010	293
Self-efficacy	2.6%	0.008	.001	913	6.8%	0.030	.023	290
Self-management	3.9%	0.011	<.001	845	2.4%	0.033	.457	293

Note. Each coefficient represents the percentage point change in a given degree completion outcome associated with an increase of one standard deviation in a given deeper learning competency measure. Results are based on the subsample of study participants who entered Grade 9 in 2009–10 or 2010–11 within 19 sampled schools, and for whom we were able to collect both high school survey and National Student Clearinghouse data. Sample sizes vary across analyses due to missing data on relevant high school survey measures.

References

Akey, T. M. (2006). *School context, student attitudes and behavior, and academic achievement: An exploratory analysis*. MDRC. https://files.eric.ed.gov/fulltext/ED489760.pdf

Buck Institute for Education. (2007). *National Survey of High School Reform and Project Based Learning*. https://www.academia.edu/5368265/National_Survey_of_PBL_and_High_School_Reform_-_Instrument

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods*, *4*(1), 62–83. https://doi.org/10.1177/109442810141004

Consortium on Chicago School Research. (2007). *Survey of Chicago Public Schools: 9th and 10th grade student edition*.

Culture of Excellence & Ethics Assessment. (2019). *Secondary level sample: Student survey items*. http://www.excellenceandethics.com/assess/2_CEEA_Student_Items_Sample_Secondary.pdf

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, *91*(2), 166–174. https://doi.org/10.1080/00223890802634290

Funk, M. J., Westreich, D., Wiesen, C., Sturmer, T., Brookhart, M. A., & Davidian, M. (2011). Doubly robust estimation of causal effects. *American Journal of Epidemiology*, *173*(7), 761–767. https://doi.org/10.1093/aje/kwq439

Hirano, K., Imbens, G. W., & Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), 1161–1189. https://scholar.harvard.edu/imbens/files/efficient_estimation_of_average_treatment_effects_using_the_estimated_propensity_score.pdf

Huang, D., Leon, S., Hodson, C., La Torre, D., Obregon, N., & Rivera, G. (2010). *Exploring the effect of afterschool participation on students' collaboration skills, oral communication skills, and self-efficacy* (CRESST Report 777). University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Learning Point Associates. (2012). *Learning Point School Climate Surveys [SCS]: High School Student Survey*. https://web.archive.org/web/20170329124051/http:/goal.learningpt.org/winss/scs/sampques.asp?survey=HS

Levenson, H. (1981). Differentiating among internality, powerful others, and chance. In H. M. Lefcourt (Ed.), *Research with the locus of control construct* (Vol. 1, pp. 15–63). Academic Press.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403–425. https://doi.org/10.1037/1082-989X.9.4.403

National Survey of Student Engagement. (2011). *NSSE 2011 U.S. English Version*. Trustees of Indiana University.

New York City. (2011). *The NYC School Survey: 2011 student survey form*.

Pace, C. R., & Kuh, G. D. (1998). *College Student Experience Questionnaire (CSEQ)*. Indiana University. http://cseq.indiana.edu/pdf/cseq_whole.pdf

Pintrich, R. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Ridgeway, G., McCaffrey, D., Morral, A., Cefalu, M., Burgette, L., Pane, J., & Griffin, B. A. (2013). *Twang: Toolkit for weighting and analysis of nonequivalent groups* (R package version 1.3-11). https://cran.r-project.org/web/packages/twang/vignettes/twang.pdf

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Tinio, M. F. O. (2009). Academic engagement scale for grade school students. *The assessment handbook*, *2*(1), 64–75.

Walker, S. L., & Fraser, B. J. (2005). Development and validation of an instrument for assessing distance education learning environments in higher education: The Distance Education Learning Environments Survey (DELES). *Learning Environments Research*, *8*(3), 289–308. https://doi.org/10.1007/s10984-005-1568-3

What Works Clearinghouse. (2020). *What Works Clearinghouse procedures handbook* (Version 4.1). U.S. Department of Education, Institute of Education Sciences. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Procedures-Handbook-v4-1-508.pdf

Xue, G., & Sun, X. (2011). Construction and validation of self-management scale for undergraduate students. *Creative Education*, *2*(2), 142–147. https://doi.org/10.4236/ce.2011.22020

Yen, W. M. (1986). The choice of scale for educational measurement: An IRT perspective. *Journal of Educational Measurement*, *23*(4), 299–325. https://www.jstor.org/stable/1434551

About the American Institutes for Research

Established in 1946, with headquarters in Arlington, Virginia, the American Institutes for Research® (AIR®) is a nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance to solve some of the most urgent challenges in the U.S. and around the world. We advance evidence in the areas of education, health, the workforce, human services, and international development to create a better, more equitable world. The AIR family of organizations now includes IMPAQ, Maher & Maher, and Kimetrica. For more information, visit AIR.ORG.

AIR® Headquarters
1400 Crystal Drive, 10th Floor
Arlington, VA 22202-3289
+1.202.403.5000 | **AIR.ORG**