



AmeriCorps Promise Fellow Program Minnesota Alliance With Youth Impact Evaluation Report, 2019

Prepared by
The **Improve** Group

Contents

Introduction	3
About The Alliance's AmeriCorps Promise Fellow Program	3
Methodology	4
Regression Discontinuity Design	5
Data Sources	6
Data Cleaning	7
Findings	8
Attendance	8
Engagement	12
Work Completion	15
Appendix A. Expanded Methodology	18

Photos by Uche Iroegbu, Minnesota Alliance With Youth, and ServeMN



Introduction

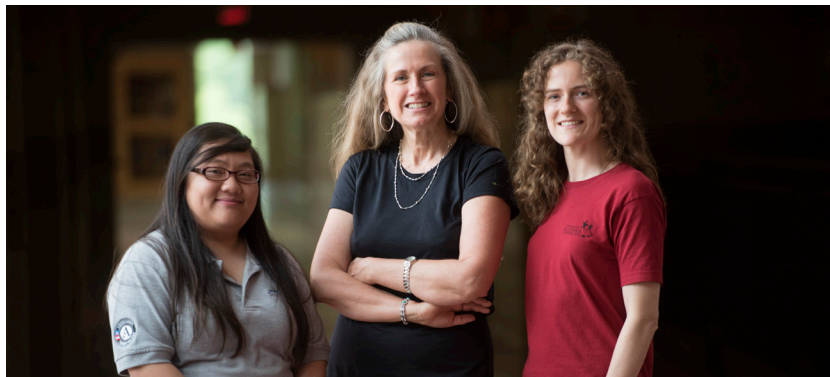
Nearly one in five high school students in Minnesota does not graduate on time. Dropping out of high school can have severe, lifelong implications, including a higher likelihood of unemployment, being a recipient eligible for public program benefits, experiencing poor health outcomes, serving time in jail, and having an expected annual salary about one-third of high school graduates (Rumberger, 2011, US Census Bureau). High school graduation also has broader societal and economic implications. For example, a 2013 study from the Alliance for Excellent Education found that the state of Minnesota would see an increase of \$82 million in annual gross state product if the graduation was increased from the current 83.2 percent to 90 percent.

Research also suggests that many of the causes of dropping out of high school are preventable, and that many risk factors are identifiable even before students enter high school. Most notably, early warning indicators of disengagement from school, including attendance, behavior, and course performance, can identify students during middle school as having a 75 percent or more likelihood of dropping out (Allenworth & Easton, 2007; Balfanz & Byrnes, 2010; Balfanz, 2009; Balfanz, Herzog & MacIver, 2007; Heppen & Therriault, 2008; Neild & Balfanz, 2006). Students leaving high school also often cite lack of motivation, boredom, an unchallenging atmosphere, or an overall lack of engagement in school as reasons for dropping out (National High School Center, 2007). Accordingly, the Minnesota Alliance With Youth's (the Alliance) AmeriCorps Promise Fellow program has been designed to specifically address these early indications of disengagement from school.

About The Alliance's AmeriCorps Promise Fellow Program

With the goal of increasing graduation rates, the Alliance hosts 220 AmeriCorps members through the Promise Fellow program. Members are placed in schools and community sites across the state and work directly with up to 30 students in grades 6 through 12 who are off-track to graduate. Members work with teams including teachers, school administrators, parents, school social workers, and community partners to identify students with low school attendance, low work completion, and/or low school engagement. Qualifying students who are placed on the members' Focus Lists participate in interventions that fall into three primary categories:

- Caring Adults, including informal mentoring, assistance with work completion, and/or check-ins throughout the day to establish more positive and supportive relationships at school.
- Service and Service-Learning, where students serve in a mentor role, volunteer, or engage in a service-learning project.
- Out-of-School Supports, such as joining after-school clubs or engaging in after-school tutoring.



Methodology

Throughout their time working with students on their Focus Lists, Promise Fellows track key indicators of progress, including demographic information about the students served and the dates and total number of minutes students are involved in the program. They also track updates on student attendance either three or four times per year, data on work completion, and ratings on student engagement from the University of Minnesota’s Student Engagement Instrument.¹ The evaluation used these indicators of progress because all can be measured within a single academic school year, which aligns with the program’s model. These same indicators are collected for students before they first begin working with Promise Fellows, to be used as baseline information, and throughout the year to measure change over time.

About the evaluation

This report was prepared by The Improve Group, a full-service evaluation, research, and strategic consulting firm based in St. Paul, Minnesota, in partnership with Research Consultant Dr. Christopher Desjardins, Ph.D. The Improve Group conducts rigorous studies to help organizations make the most of information, navigate complexity, and ensure their investments lead to meaningful, sustained impact. The Alliance and The Improve Group have been partners in evaluation since 2016.

1 <http://checkandconnect.umn.edu/sei/default.html>

Regression Discontinuity Design

To evaluate the progress of the Promise Fellow program, the Alliance and evaluation partners designed a study for the 2018-2019 school year answering three primary evaluation questions:

- 1. To what extent do students served by the Alliance demonstrate greater increases in monthly attendance relative to other students who do not qualify to be served, but are close to the cut-off?**
- 2. To what extent do students served by the Alliance demonstrate greater increases in pre-post engagement relative to other students who do not qualify to be served, but are close to the cut-off?**
- 3. To what extent do students served by the Alliance demonstrate greater increases in monthly work completion relative to other students who qualify to be served, but are close to the cut-off?**

To answer these evaluation questions, evaluators employed a regression discontinuity design (RDD). RDD is commonly used in studies that compare pre- and post-test data for two groups: one that received a treatment and one that did not. What distinguishes RDD from a randomized control or quasi-experimental design is that assignment to the treatment or control group is not random, but instead based on a characteristic that creates a qualification for treatment (Reardon, Robinson 2010).

In this case, evaluators used a two-criteria regression discontinuity because there are two criteria for being served by a Promise Fellow: attendance and engagement. Students qualified to work with the Promise Fellow program if they:

1. Had an overall attendance rate below 90 percent; or
2. Had an engagement score from the SEI less than or equal to 100.

Students who qualified to be served by Promise Fellows based on attendance and/or engagement made up the treatment group (referred to throughout this report as “Focus List students”). Students who did not qualify but were close to the attendance and/or engagement cutoffs made up the Control Group. Therefore, the two-criteria regression discontinuity model incorporated the two eligibility criteria, as well as a dummy variable for whether students were in the Control Group or on the Focus List, when measuring program effect.

Promise Fellows collected data on attendance, student engagement, and work completion (a third outcome expected to improve by working with Promise Fellows, but which was not used to determine program eligibility) for all Focus List and Control Group students. This data collection occurred during an initial baseline period (prior to service) and throughout the year during benchmark periods.

Data Sources

Each evaluation question was informed by a student-level data source shared by schools with Promise Fellows.

1. To what extent do students served by the Alliance demonstrate greater increases in monthly attendance relative to other students who do not qualify to be served, but are close to the cut-off?

School attendance records informed this evaluation question. Promise Fellows obtained records on the number of days students were absent from or tardy to school during an initial period prior to working together (baseline data), as well as in quarterly increments throughout the duration of the school year (benchmark data). Students qualified for the program if they attended school on fewer than 90 percent of days during the baseline period. Teachers shared data for these Focus List students and for students in the Control Group who did not qualify to be served.

2. To what extent do students served by the Alliance demonstrate greater increases in pre-post engagement relative to other students who do not qualify to be served, but are close to the cut-off?

Student engagement was measured using the University of Minnesota's Student Engagement Instrument (SEI). Students self-reported their ratings to 35 questions using a four-point scale, with possible total scores between 35 and 140. A higher score indicated greater student engagement. The SEI was administered to students during one baseline period and during several benchmark periods throughout the school year. Focus List students qualified if they scored 100 or less on the SEI during the baseline period; Control Group students all scored higher than 100 at the baseline.

3. To what extent do students served by the Alliance demonstrate greater increases in monthly work completion relative to other students who qualify to be served, but are close to the cut-off?

Teacher input informed work completion data. Each month, teachers shared a summary of the percentage of assignments students submitted. Just as with attendance and engagement data, work completion scores were also shared from an initial baseline period, prior to engagement with Promise Fellows and throughout the remainder of the school year for benchmarking. Work completion was not a criterion for eligibility in the program. All Focus List students included in the analysis of this evaluation question qualified based on attendance and/or engagement. Work completion records were collected for both Focus List students and Control Group students who did not qualify for programming.



Data Cleaning

Initial data cleaning for the RDD was done to remove students with data that either fell outside of plausible values or fell on the wrong side of the discontinuity point for attendance or engagement. For example, Focus List students were excluded from attendance analysis if they had a baseline attendance rate greater than or equal to 90 percent, and Control Group students were excluded if they had attendance rates less than 90 percent. Similarly, Focus List students were excluded from the engagement analyses if they had an initial engagement score greater than 100, as were Control Group students with engagement of less than or equal to 100. For analysis of work completion, Focus List students who had both attendance greater than 90 percent and Student Engagement of greater than 100 were excluded. Control Group students with attendance of less than 90 percent and engagement of less than or equal to 100 were also excluded.

In addition to the RDD, evaluators analyzed progress on outcomes from Focus List students without a comparison to the Control Group. They also conducted a series of sub-analyses to explore potential progress within different student groups. The following includes the results of these analyses; an expanded methodology and full coefficient outputs are available in Appendix A.



Findings

While the Regression Discontinuity Design (RDD) found the Promise Fellow program had no effect on attendance, work completion, or student engagement, Focus List students did demonstrate improvements in all three of these key outcome areas, even as students in the Control Group showed declines. This, and other preliminary evidence of success, are discussed below.

Attendance

To what extent do students served by the Alliance demonstrate greater increases in monthly attendance relative to other students who do not qualify to be served, but are close to the cut-off?

Schools shared attendance records with Promise Fellows for all Focus List and Control Group students. Students who qualified for the program based on attendance were in school on fewer than 90 percent of days during the initial baseline period. The program used the 90 percent attendance threshold because it meets the Minnesota Department of Education's definition of chronic absenteeism under the Every Student Succeeds Act.¹ Focus List students attended school on fewer than 90 percent of days, and Control Group students attended 90 percent or more of school days. Data about tardiness was also collected and used in additional analyses (described below) to further inform this evaluation question.

As shown in Table 1, Focus List students and students in the Control Group were similar demographically, except that Control Group students were less likely to be eligible for free or reduced-price lunch and more likely to be white and Asian.

² <https://education.mn.gov/MDE/dse/ESSA/Imp/MDE073339>

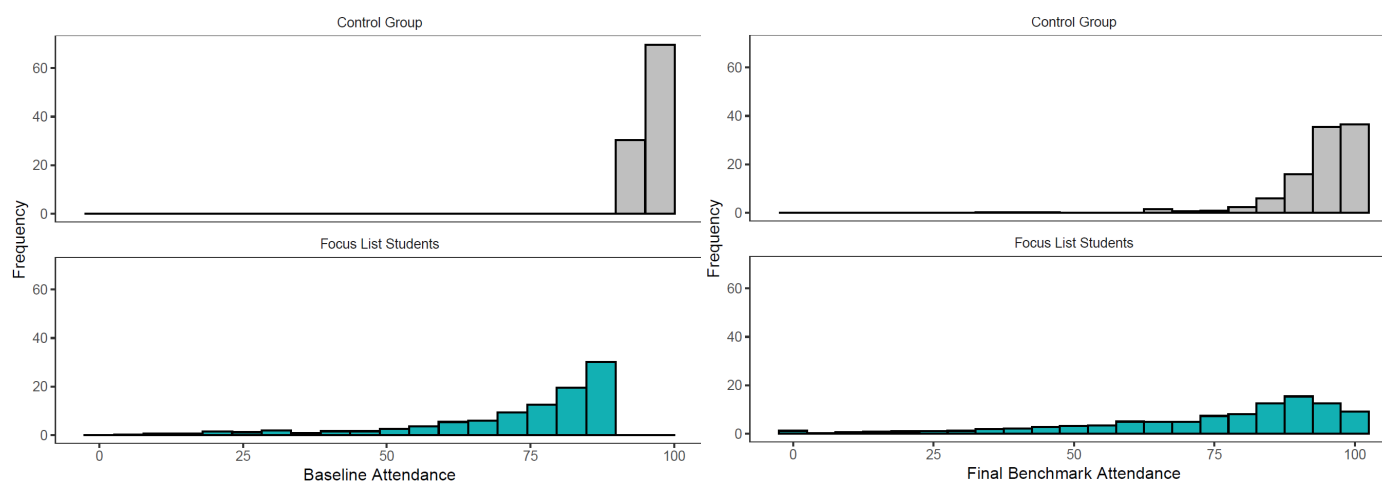
Table 1. Student Demographics for Attendance Analyses

Variable	Control Group	Focus List
Group Size	378	1,900
Gender		
Male	52%	50%
Female	48%	50%
Race/Ethnicity		
African American	38%	46%
White	35%	25%
Hispanic/Latinx	10%	17%
Asian	9%	3%
American Indian	1%	6%
Native Hawaiian	2%	0%
Additional Demographic Characteristics		
English Language Learner	9%	10%
Disabled	5%	9%
Free/Reduced Price Lunch	15%	20%
Parent	1%	3%
Homeless/Highly Mobile	1%	3%

To evaluate attendance data for program effect, evaluators used a Regression Discontinuity Design, with the discontinuity set at the 90 percent attendance threshold during the baseline data collection period; while Control Group students were all above but near that threshold, all Focus List students were below it, with some students having very low attendance (including about 11 percent who attended school on fewer than 50 percent of days). Figure 1 below shows the distribution of attendance for Control Group and Focus List students both during the baseline period before programming and during the final benchmark period.

As seen in Figure 1, the distribution of students between the baseline and benchmark periods shifted for both the Control Group and Focus List students, but in opposite directions; Focus List students shifted to the left, with more students attending at a higher rate, while Control Group students shifted to the right (more students attending at a lower rate). **Whereas no Focus List students exceeded the 90 percent attendance threshold during the baseline period, 30 percent did during the final benchmark.** Control Group students shifted from 100 percent exceeding the 90 percent attendance threshold during the baseline period to 84 percent exceeding that threshold during the final benchmark.

Figure 1. Baseline and Final Attendance Distributions for Control Group and Focus List Students



Whereas the shift seen between the baseline and benchmark attendance periods is in the expected direction, running the RDD statistical model did not indicate a positive program effect, while controlling for program site. As shown in Table 2, there was a statistically significant difference between the Control Group and Focus List students. The Focus List students had a 4 percent lower expected attendance rate relative to Control Group students. This suggests that Focus List students did not improve relative to students in the Control Group.

Table 2. RDD Coefficients for the Attendance Model

Parameter	Estimate	SE	df	t	p
Expected Attendance for Control Group	0.878	0.015	226.734	58.878	0.000
Program Effect	-0.043	0.012	2,025.463	-3.535	0.000

Following the initial RDD, evaluators conducted a supplemental RDD analysis, this time only with the students closest to the 90 percent threshold, (those with a baseline attendance between 85 percent and 95 percent). As described previously, many students on the Focus List had very low attendance (including about 11 percent who attended school on fewer than 50 percent of days), and many students in the Control Group had perfect attendance. In theory, students with very low attendance would not be experiencing the Promise Fellow program as designed, as they simply are not present enough to participate. On the other hand, perfect attendance can be problematic for analyzing program effect, too, since there is no space for improvement for these students. With these tighter parameters, this secondary analysis included 509 Focus List students and 103 Control Group students.

As shown in Table 3, when only including students closer to the 90 percent discontinuity, the statistical model did not indicate a positive program effect. While controlling for program site and baseline attendance, the Focus List students had a 3.8 percent lower expected attendance rate relative to Control Group students. However, the p-value for being on the Focus List exceeded the .05 threshold for statistical significance, indicating that the difference between Focus List students and Control Group students was no longer significant, so the expected attendance rate did not significantly differ between the Control Group and the Focus List students.

Table 3. RDD Coefficients for the Attendance Model Only Including Students Near the Discontinuity Point

Parameter	Estimate	SE	df	t	p
Expected Attendance for Control Group	0.896	0.017	308.158	52.521	0.000
Program Effect	-0.038	0.022	600.328	-1.779	0.076

Focus List students increased attendance, while students in the Control Group decreased attendance.

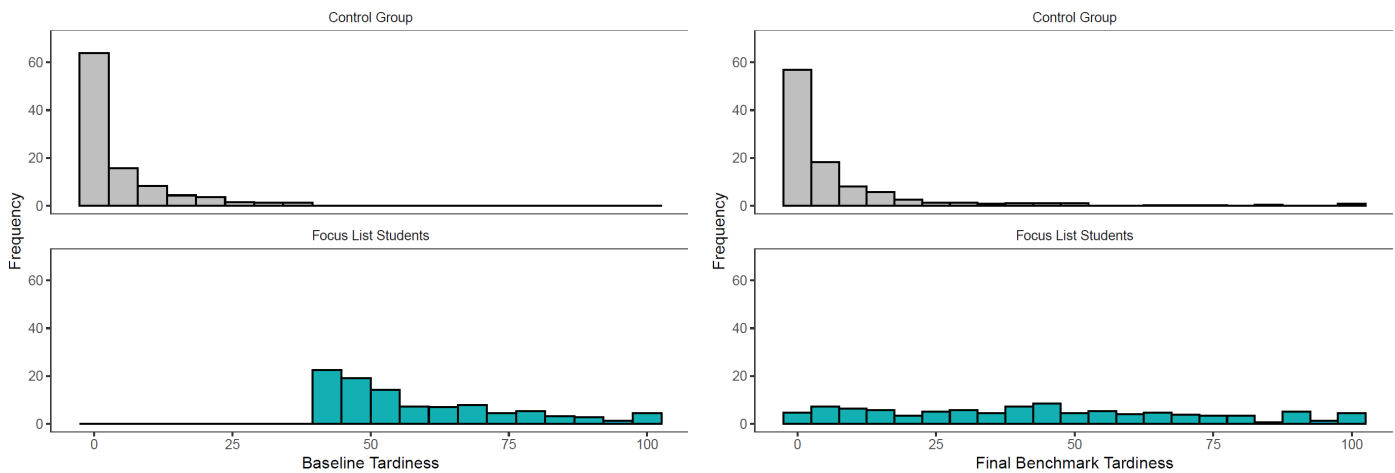
Although the RDD analysis showed no effect of the Promise Fellows on overall attendance, other analyses did show progress in the intended direction. Focus List students experienced an average attendance increase from 73 percent attendance at baseline to 75 percent attendance at their final benchmark. In total, this expected increase in attendance, after controlling for site, is equivalent to attending school 3.4 additional days in a 180-day school year. While this increase was not statistically significant when controlling for site (p=0.10), it is an improvement over the baseline.

Nearly three in five Focus List students, or 59 percent, increased their individual attendance rates between the baseline and final benchmark periods. By the final benchmark, 30 percent of Focus List students had a final benchmark attendance rate greater than 90 percent.

Focus List Students decreased tardiness by nearly one-fifth between the baseline and final benchmark periods.

While the primary evaluation question was developed to measure change in overall attendance, schools also shared data on Focus List and Control Group tardiness with Promise Fellows. To explore for potential program effects on tardiness, evaluators conducted an RDD analysis on attendance. The analysis was like the one that was run for overall attendance but with a baseline discontinuity point at tardy for 40 percent of school days (as opposed to the 90 percent discontinuity for attendance). Although no effect was shown in the RDD analysis, Focus List students had a substantial decrease in tardiness between the baseline and final benchmark periods, while Control Group students experienced a slight increase during that same time period. Overall, Focus List students reduced average tardiness from 59 percent during the baseline period to 44 percent during the final benchmark, while Control Group students' tardiness increased from five percent to eight percent. Figure 2 illustrates the distribution of Focus List and Control Group students during the baseline and final benchmark periods; it also shows the big shift Focus List students saw toward being tardy less often.

Figure 2. Baseline and Final Tardiness Distributions for Control Group and Focus List Students



While the RDD did not indicate a program effect, analysis of just the Focus List students, without the Control Group, did reveal preliminary evidence for program impact. First, most Focus List students saw reductions in their individual tardiness rates: 69 percent of Focus List students decreased their individual tardiness between the benchmark and baseline periods. When controlling for site, Focus List students were expected to decrease tardiness by nearly 20 percent, the equivalent of coming to school on time 35 additional days during a 180-day school year.

Engagement

To what extent do students served by the Alliance demonstrate greater increases in pre-post engagement relative to other students who do not qualify to be served, but are close to the cut-off?

Student engagement is measured using the University of Minnesota’s Student Engagement Instrument (SEI). Students self-report their personal ratings to 35 questions, each on a scale from one to four (with one being least engaged and four being most). To evaluate for program effect, evaluators used a Regression Discontinuity Design, which used data from both Focus List students and the Control Group, and set the discontinuity point at 100 on the SEI. This score was the Promise Fellow eligibility criteria for students who qualified based on engagement.

Focus List students and Control Group students took the SEI at the start of the school year (prior to Focus List students beginning to work with Promise Fellows) to serve as a baseline point. They also took the SEI throughout the school year to provide benchmark data. Table 4 below includes the demographic breakdown of the Focus List and Control Group students. The two groups were largely similar, with the control group more likely to be white, and less likely to be Hispanic/Latinx, African American, and eligible for free or reduced-price lunch.

Table 4. Student Demographics for Engagement Analysis

Variable	Control Group	Focus List
Group Size	338	1,600
Gender		
Male	50%	51%
Female	50%	49%
Race/Ethnicity		
African American	43%	36%
White	32%	37%
Hispanic/Latinx	11%	17%
Asian	5%	2%
American Indian	1%	5%
Native Hawaiian	2%	0%
Additional Demographic Characteristics		
English Language Learner	7%	10%
Disabled	5%	8%
Free/Reduced Price Lunch	17%	22%
Parent	0%	2%
Homeless/Highly Mobile	1%	2%

Figure 3 below illustrates the distribution of student engagement for Control Group and Focus List students both during the baseline and final benchmark reporting periods. As shown, the discontinuity point was 100 on the SEI; while Control Group students were all above that threshold at the baseline, all Focus List students were below it. For the most part, Control Group students were near the discontinuity threshold; however, some Focus List students were quite far below that point, including about 7 percent who scored lower than 75 during the baseline period.

Figure 3. Baseline and Final Student Engagement Distributions for Control Group and Focus List Students

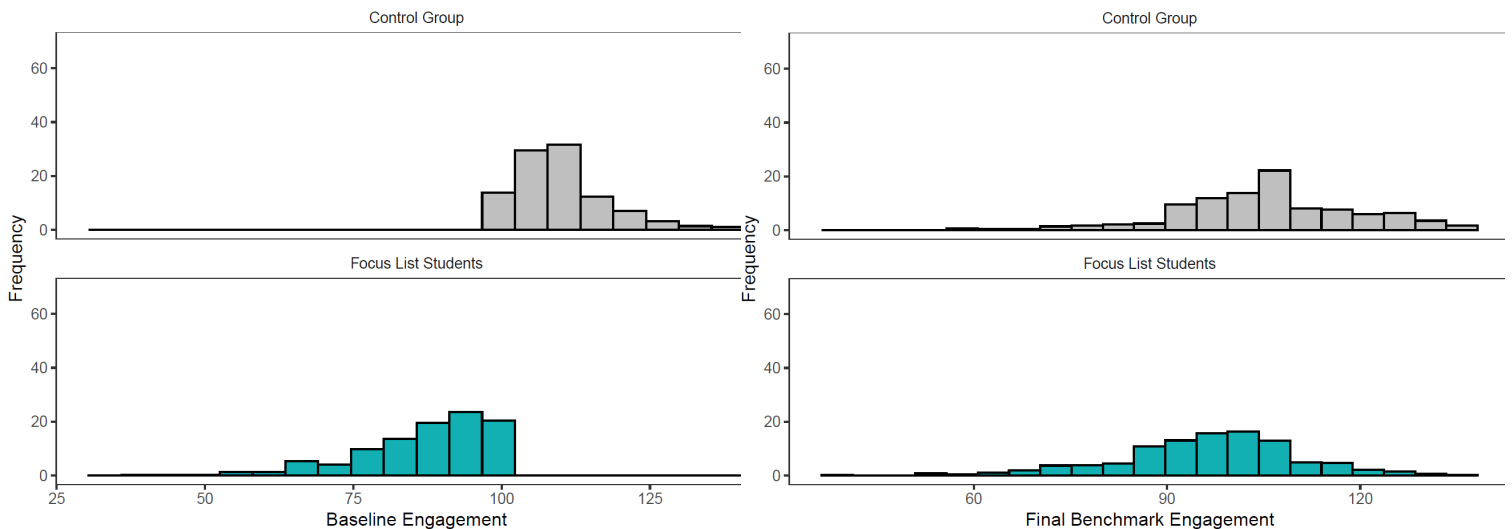


Figure 3 also shows the shifts in the distribution of students between the baseline and benchmark periods; Focus List students shifted to the left, with more students scoring higher on the SEI during the final benchmark period, while Control Group students shifted to the right, with more students scoring lower on the SEI. Whereas no Focus List students exceeded the 100 SEI score threshold during the baseline period, 40 percent did during the final benchmark. Control Group students shifted from 100 percent exceeding the 100 SEI threshold during the baseline period to 68 percent during the final benchmark.

Whereas the shift seen between the baseline and benchmark SEI reporting periods is vast and in the direction that would be expected from the program participants, running the RDD statistical model did not indicate a positive program effect. The model controlled for site, and, as shown in Table 5, estimated that the expected engagement for Control Group students had a two-point lower expected SEI score relative to Control Group students. However, because the p-value is greater than .05, there was not a statistically significant difference between the Control Group and Focus List students.

Table 5. RDD Coefficients for the Student Engagement Model

Parameter	Estimate	SE	df	t	p
Expected Engagement for Control Group	104.915	1.181	198.099	88.836	0.000
Program Effect	-2.072	1.142	1,422.887	-1.814	0.070

As with Attendance, following the initial RDD, evaluators conducted a supplemental RDD analysis only including students who were close to the 100 SEI discontinuity point. In this case, the analysis included only students who had scored between 86 and 114 on the baseline SEI, which is equal to one standard deviation in either direction. As with attendance, the narrower analysis was conducted because students with very high or very low engagement can mask program effect for students who score nearer to those who are typically involved. With the narrower parameters, the analysis included 730 Focus List students and 220 students from the Control Group.

As is shown in Table 6, when only including students closer to the 100 SEI discontinuity, the effect size for being on the Focus List is still negative, though much smaller (-0.6, compared to -2.0 when all students were included), and the difference between Focus List students and Control Group students is still not statistically significant.

Table 6. RDD Coefficients for the Student Engagement Model Only Including Students Near the Discontinuity Point

Parameter	Estimate	SE	df	t	p
Expected Engagement for Control Group	102.868	1.202	335.750	85.609	0.000
Program Effect	-0.641	1.530	945.823	-0.419	0.675



Focus List students significantly improved their engagement over the course of the school year.

While the primary evaluation question for measuring student engagement was developed to use a Regression Discontinuity Design, which compared Focus List students to those in the Control Group, evaluators also conducted an analysis to measure the change that happened within the Focus List group independently of the Control Group. For Focus List students, the average baseline engagement score was 87, which improved by nine points to 96 during the final baseline period. When looking at individual progress, nearly 74 percent of students on the Focus List increased their engagement between the baseline and final benchmark. While no Focus List students were above the 100 SEI cutoff at the first baseline, 40 percent of students scored 100 or higher on the SEI in the final benchmark period. In total, the expected increase in engagement for Focus List students, after controlling for site, is 9.7 points out of 140 on the SEI, which is statistically significant ($p < .001$).

Work Completion

To what extent do students served by the Alliance demonstrate greater increases in monthly work completion relative to other students who qualify to be served, but are close to the cut-off?

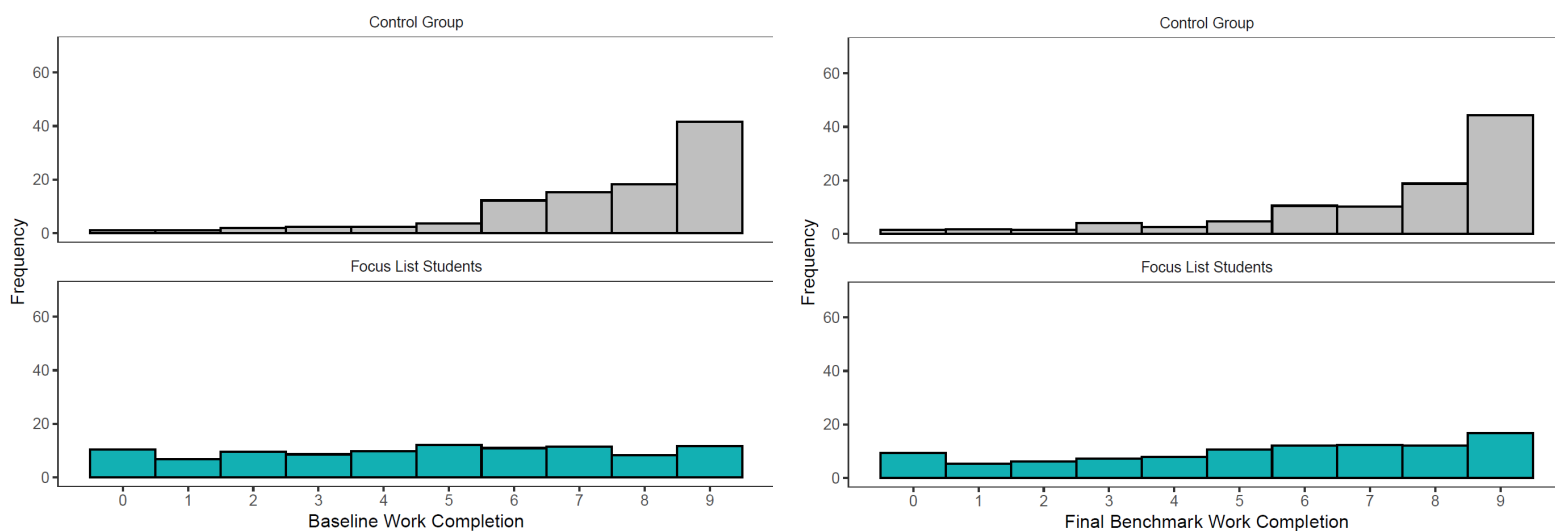
Teachers shared with Promise Fellows monthly information about work completion for Focus List and Control Group students; the information was then converted to a 10-point scale. Unlike for attendance and engagement, evaluators did not use a Regression Discontinuity Design (RDD) to analyze change between students on the Focus List and those in the Control Group. Attendance and engagement are used as eligibility criteria for participating in the Promise Fellow program (students served either start the year with attendance lower than 90 percent and/or score lower than 100 on the Student Engagement Instrument). Therefore, these two outcomes had built-in discontinuity points around which to design the RDD. Furthermore, students were only included in the analysis for attendance or engagement if they qualified based on that outcome. Work completion is a third outcome expected to improve from being on a Promise Fellow's Focus List; however, it was not a criterion for being included on the list. Therefore, rather than doing RDD analysis, evaluators fit a linear mixed effects model that included all students, whether they qualified based on engagement, attendance, or both. Table 7 lists the demographics of students on the Focus List and those in the Control Group for the work completion analysis. The two groups were generally similar, with the control group more likely to be white and Asian, and less likely to be African American, Hispanic/Latinx, and eligible for free and reduced-price lunch.

Table 7. Student Demographics for Work Completion Analyses

Variable	Control Group	Focus List
Group Size	456	2,773
Gender		
Male	53%	52%
Female	47%	48%
Race/Ethnicity		
African American	35%	43%
White	38%	30%
Hispanic/Latinx	11%	17%
Asian	8%	3%
American Indian	1%	5%
Native Hawaiian	2%	0%
Additional Demographic Characteristics		
English Language Learner	9%	10%
Disabled	5%	8%
Free/Reduced Price Lunch	14%	21%
Parent	0%	2%
Homeless/Highly Mobile	1%	2%

Like attendance and student engagement, evaluators used benchmark and baseline work completion data for both Focus List students and the Control Group. The distribution of each is displayed in Figure 4. As shown, Control Group students generally had higher average work completion, and Focus List students were mostly evenly distributed. This pattern was the same during the final benchmark period, suggesting that work completion did not change over the course of the school year.

Figure 4. Baseline and Final Work Completion Distributions for Control Group and Focus List Students



The statistical model confirms that there was no program effect on work completion; as shown in Table 8, there is a statistically significant difference between Focus List students and the Control Group ($p = 0.000$). The estimate for the effect of being on the Focus List is negative.

Table 7. Work Completion Coefficients Table

Parameter	Estimate	SE	df	t	p
Expected Work Completion for Control Group	3.888	0.224	621.096	17.362	0.000
Program Effect	-0.347	0.157	2,277.956	-2.206	0.028



Focus List students significantly increased work completion over the course of the year, while Control Group students' work completion decreased slightly.

Evaluators explored Focus List students' work completion data without comparison to the Control Group and found significant improvements between the initial baseline and final benchmark periods. Focus List students' average work completion at the baseline was 4.8 out of a possible 10; by the final benchmark, the average had improved to 5.3. In total, 45 percent of Focus List students increased their work completion between the baseline and final benchmark.

Evaluators analyzed Focus List student data for preliminary evidence of program effect and found that, when controlling for site, the increase in work completion between baseline and the final benchmark for these students was statistically significant ($p < 0.001$). In total, Focus List students are expected to complete nearly 6 percent more homework between the baseline and final benchmark.

Appendix A. Expanded Methodology

Data Cleaning

Focus List participants were excluded from the main attendance analyses if they had baseline program attendance rates greater than or equal to 90 percent, and Control Group participants were excluded if they had attendance rates less than 90 percent. In addition, all participants who had attendance rates of 0 were excluded, as these values seemed implausible. For the tardiness analyses, evaluators included only Focus List participants with tardiness rates above 40 percent and control participants less than 40 percent.

Similarly, Focus List participants were excluded from the engagement analyses if they had an initial engagement score greater than 100, and Control Group participants were excluded if they had an engagement score less than or equal to 100.

Finally, participants from the work completion analyses were excluded if they were a Focus List participant who had an attendance rate that was greater than or equal to 90 percent and an initial engagement score greater than 100, or if they were Control Group students who had an attendance rate less than 90 percent and an engagement score less than or equal to 100.

Statistical Model

Evaluators fit a linear mixed effects model to answer the three principal evaluation questions. This type of model was selected because it allows evaluators to examine repeated measures over time and also accounts for the hierarchical nature of this data (in this case, individual student data nested within site), which is important because individuals at the same site have more shared variance and are more likely to have similar outcomes than individuals at other sites.

For each outcome (attendance, student engagement, and work completion) separately, evaluators fit the following model (or very similar models that always controlled for site, unless explicitly mentioned):

$$\text{Final}_{ij} = \beta_0 + b_{0j} + \beta_1 \text{Focus List}_{ij} + \beta_2 \text{Base}_{ij} + e_{ij}$$

The model states that participant i at site j 's score on the final benchmark outcome (Final) was a function of a grand intercept (β_0), a random intercept associated with the site (b_{0j}), whether they participated in the Promise Fellow program or were a control participant (Focus List_{ij}, dummy coded such that 1 if they were a program participant and 0 if they were in the Control Group), their score on the baseline outcome (Base_{ij}) and random error (e_{ij}). Evaluators assumed that $b_{0j} \sim N(0, \sigma_{b_0}^2)$, $e_{ij} \sim N(0, \sigma_e^2)$ and that $\text{cov}(b_{0j}, e_{ij}) = 0$. For the work completion model, evaluators included both baseline attendance and baseline engagement scores in addition to baseline work completion as these were used for the experimental design.

In addition, evaluators examined whether there was a multiplicative effect (e.g., that the slopes of the baseline outcome on the final benchmark outcome were not parallel for Focus List and control participants) and fit the following model,

$$\text{Post}_{ij} = \beta_0 + b_{0j} + \beta_1 \text{Focus List}_{ij} + \beta_2 \text{Pre}_{ij} + \beta_3 \text{Focus List}_{ij} : \text{Pre}_{ij} + e_{ij}$$

The statistical modeling was performed in R using the lme4 package.

Subgroup model output

Table 1. Coefficients table for the attendance subgroup models.

Model	Parameter	Estimate	SE	df	t	p
White	Intercept	0.830	0.013	120.764	62.310	0.000
	White	0.016	0.012	1,683.581	1.378	0.168
	Baseline Attendance	0.467	0.028	1,700.109	16.844	0.000
Hispanic	Intercept	0.835	0.013	114.262	63.593	0.000
	Hispanic	-0.000	0.013	1,696.152	-0.007	0.995
	Baseline Attendance	0.468	0.028	1,700.541	16.850	0.000
Male	Intercept	0.844	0.014	136.021	61.157	0.000
	Male	-0.019	0.009	1,677.576	-2.034	0.042
	Baseline Attendance	0.469	0.028	1,699.599	16.893	0.000
Community Site	Intercept	0.830	0.014	102.425	61.100	0.000
	Community Site	0.047	0.040	106.690	1.182	0.240
	Baseline Attendance	0.468	0.028	1,700.097	16.842	0.000

Table 2. Coefficients table for the tardiness subgroup models.

Model	Parameter	Estimate	SE	df	t	p
White	Intercept	0.329	0.035	70.414	9.488	0.000
	White	-0.043	0.043	310.972	-0.992	0.322
	Baseline Tardiness	0.286	0.093	307.302	3.067	0.002
Hispanic	Intercept	0.315	0.034	68.046	9.228	0.000
	Hispanic	0.045	0.039	295.150	1.153	0.250
	Baseline Tardiness	0.287	0.093	307.365	3.084	0.002
Male	Intercept	0.328	0.038	98.972	8.595	0.000
	Male	-0.011	0.028	291.028	-0.376	0.707
	Baseline Tardiness	0.287	0.093	307.276	3.072	0.002
Community Site	Intercept	0.324	0.035	59.707	9.346	0.000
	Community Site	-0.052	0.120	42.704	-0.429	0.670
	Baseline Tardiness	0.290	0.094	306.552	3.105	0.002

Table 3. Coefficients table for the engagement subgroup models.

Model	Parameter	Estimate	SE	df	t	p
White	Intercept	102.874	1.059	112.410	97.153	0.000
	White	-1.584	0.894	1,111.495	-1.772	0.077
	Baseline Engagement	0.421	0.036	1,141.657	11.555	0.000
Hispanic	Intercept	102.432	1.034	105.844	99.092	0.000
	Hispanic	-0.503	1.005	1,143.862	-0.500	0.617
	Baseline Engagement	0.422	0.037	1,142.114	11.525	0.000
Male	Intercept	102.711	1.066	114.884	96.351	0.000
	Male	-0.838	0.739	1,120.434	-1.134	0.257
	Baseline Engagement	0.420	0.037	1,140.442	11.453	0.000
Community Site	Intercept	102.116	1.068	89.890	95.624	0.000
	Community Site	1.728	2.783	67.332	0.621	0.537
	Baseline Engagement	0.420	0.037	1,141.509	11.507	0.000

Table 4. Coefficients table for the work completion subgroup models.

Model	Parameter	Estimate	SE	df	t	p
White	Intercept	3.446	0.179	248.131	19.231	0.000
	White	0.346	0.130	1,815.756	2.672	0.008
	Baseline Attendance	2.002	0.352	1,887.176	5.691	0.000
	Baseline Engagement	0.013	0.004	1,936.330	3.471	0.001
	Baseline Work Completion	0.432	0.022	1,922.533	19.761	0.000
Hispanic	Intercept	3.593	0.179	243.992	20.078	0.000
	Hispanic	-0.209	0.146	1,963.227	-1.433	0.152
	Baseline Attendance	2.028	0.352	1,898.122	5.754	0.000
	Baseline Engagement	0.013	0.004	1,939.635	3.479	0.001
	Baseline Work Completion	0.434	0.022	1,925.913	19.821	0.000
Male	Intercept	3.674	0.189	296.436	19.438	0.000
	Male	-0.194	0.108	1,954.219	-1.803	0.072
	Baseline Attendance	2.067	0.353	1,896.681	5.850	0.000
	Baseline Engagement	0.013	0.004	1,938.396	3.392	0.001
	Baseline Work Completion	0.429	0.022	1,922.251	19.495	0.000
Community Site	Intercept	3.496	0.180	203.533	19.434	0.000
	Community Site	0.559	0.406	85.393	1.378	0.172
	Baseline Attendance	1.995	0.353	1,893.411	5.659	0.000
	Baseline Engagement	0.013	0.004	1,934.818	3.423	0.001
	Baseline Work Completion	0.432	0.022	1,932.295	19.717	0.000

Appendix B. Works Cited

Allenworth & Easton (2007). What Matters for Staying On-Track and Graduating in Chicago Public Schools. Consortium on Chicago School Research: University of Chicago.

Alliance for Excellent Education (2013). Saving Futures, Saving Dollars: The Impact of Education on Crime Reduction and Earnings. Washington, DC.

Balfanz, R. (2009). Putting middle grades students on the graduation path: A policy and practice brief. Westerville, OH: National Middle School Association.

Balfanz & Byrnes (2010). The Importance of Being in School: A Report on Absenteeism in the Nation's Public Schools. Johns Hopkins University.

Balfanz, Herzog & MacIver (2007). Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions. *The Journal of Educational Psychology*.

Heppen & Therriault (2008). Developing Early Warning Systems to Identify Potential high School Dropouts. The National High School Center.

National High School Center (2007). Easing the Transition to High School: Research and Best Practices Designed to Support High School Learning. The National High School Center.

Neild & Balfanz (2006). Unfulfilled Promise: The Dimensions and Characteristics of Philadelphia's Dropout Crisis, 2000-2005. The Philadelphia Youth Network.

Reardon, S.F., & Robinson, J.P. (2010). Regression discontinuity designs with multiple rating-score variables.

Rumberger, R. W. (2011). Dropping out: Why students drop out of high school and what can be done about it. Cambridge, MA, US: Harvard University Press.

US Census Bureau (2002). The Big Payoff: Educational Attainment and Synthetic Estimates of Work-Life Earnings. US Department of Commerce: Economics and Statistics Administration.