CENTER FOR SUPPORTIVE SCHOOLS

In School and On-Track to Graduate:

Key Findings from the AT&T Aspire-Funded Evaluation of the Peer Group Connection Cross-Age Peer Mentoring and High School Transition Program

In School and On-Track to Graduate:

Key Findings from the AT&T Aspire-Funded Evaluation of the Peer Group Connection Cross-Age Peer Mentoring and High School Transition Program

In 2014, AT&T contributed nearly $12.5 million to an initiative to support meeting GradNation's goal of a 90% national high school graduation rate by 2020. Over 1,100 proposals were submitted to the *AT&T Aspire initiative* competition and 30 projects were funded. The Center for Supportive Schools' (CSS) Peer Group Connection-High School (PGC-HS) was selected to be implemented over two years (2014-15 and 2015-16) as part of two distinct projects serving a total of 32 schools: a project to support 28 high schools in high-need urban communities in Baltimore, Maryland and New York City; and a project to support four high schools in the high-need rural community of Sampson County, North Carolina. These two projects received approximately $1.1 million in Aspire funding.

> The PGC-HS program is moving the needle in academic outcomes.
>
> ~ Westat

PGC-HS is an evidence-based, cross-age peer mentoring program that supports and eases the critical period of students' transition from middle to high school. Older students, known as peer leaders, are enrolled in a daily, for-credit, yearlong leadership course taught by school faculty during regular school hours. Each week, pairs of peer leaders meet with small groups of younger students in outreach sessions designed to strengthen relationships and build social and emotional skills.

AT&T contracted with Westat Inc., an independent evaluator, to develop a comprehensive data-driven approach to estimate the effect of the 30 AT&T Aspire-funded programs on high school student outcomes over two years. Westat developed a report[i] that included key findings from the study of PGC-HS:

- 9th grade students and peer leaders who participated in PGC-HS attended school more often than non-participants.
- Increased attendance rates were sustained a full year after students completed PGC-HS.
- 9th grade students who participated in PGC-HS were on-track to graduate on time more so than were non-participants, a finding that was also sustained into 10th grade, one year after PGC-HS participation.
- 9th grade students and peer leaders who participated in PGC-HS exhibited a higher grade point average (GPA[ii]) than non-participants.
- The more time students spent participating in PGC-HS, the better their academic outcomes.

9th grade students and peer leaders who participated in PGC-HS attended school more often than non-participants.

- In Year 1 of the study, PGC-HS participants in Baltimore and New York City exhibited a **3.4% higher attendance rate** than non-participants. Based on a 180-day school year, this means that PGC-HS participants attended **6.1 more days of school** than non-participants.

- In Year 1 of the study, PGC-HS participants in North Carolina exhibited a **3.5% higher attendance rate** than non-participants. Based on a 180-day school year, this means that PGC-HS participants attended **6.2 more days of school** than non-participants.

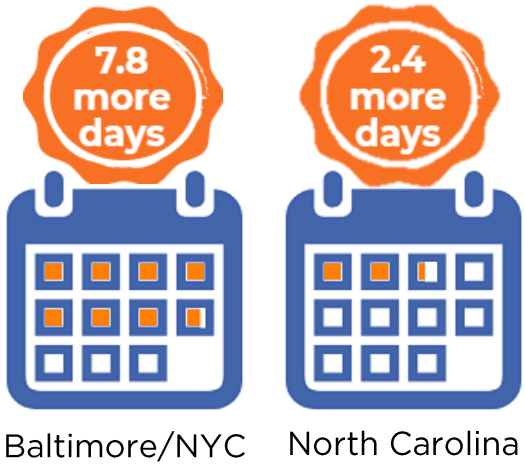
PGC participants attended more days of school than non-PGC participants

Particularly impressive is the finding for both programs, that Year 1 students see sustained advantages for school attendance in Year 2 even when no longer attending PGC-HS.

~ Westat

Increased attendance rates were sustained a full year after students completed PGC-HS.

One year later, PGC participants attended more days of school than non-PGC participants



- In Baltimore and New York City, Year 1 PGC-HS students assessed at the end of Year 2 (the year after completing PGC-HS) exhibited a **4.3% higher attendance rate** than non-participants. Based on a 180-day school year, this means that PGC-HS participants attended **7.8 more days of school** than non-participants.

- In North Carolina, Year 1 PGC-HS students assessed at the end of Year 2 (the year after completing PGC-HS) exhibited a **1.3% higher attendance rate** than non-participants. Based on a 180-day school year, this means that PGC-HS participants attended **2.4 more days of school** than non-participants.

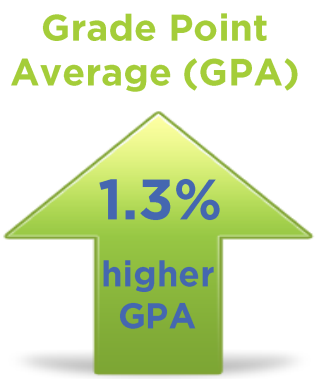
> The fact that CSS successfully demonstrated impacts in both regions [Baltimore/New York City and North Carolina] is impressive. This finding suggests that PGC-HS might scale successfully to other contexts as well; a hypothesis which would need to be validated through future research.
>
> ~ Westat

9th grade students who participated in PGC-HS were on-track to graduate on time more so than were non-participants, a finding that was also sustained into 10th grade, one year after PGC-HS participation.

- The *On-track to On-time Graduation Indicator (OTI)* was developed by researchers to predict whether a student was on-track for on-time graduation. It takes into consideration for each student the number of credits earned and the number of core courses failed during a school year.
- Results among Year 1 PGC-HS participants in Baltimore and New York City[iii] assessed at the end of Year 2 showed a sustained moderate effect size difference (.59) for OTI.

9th grade students and peer leaders who participated in PGC-HS exhibited a higher grade point average (GPA) than non-participants.

In Year 2, PGC-HS participants in Baltimore and New York City[iv] exhibited a **1.3% higher mean GPA** than matched comparison students. This impact finding was statistically significant.



The more time students spent participating in PGC-HS, the better their academic outcomes.

Students were divided into two groups based on the median number of days of participation. Students who participated more than the median number of days were considered *high-intensity* students and those who participated less than or equal to the median number of days were considered *low-intensity* students. Program impacts were compared between the *high-intensity* and *low-intensity* subgroups.

- Across all regions, students participating in the high-intensity PGC-HS subgroup (top 50% of contact days) had statistically significantly higher outcomes in OTI, GPA, and school attendance than students in the low-intensity subgroup (lower 50% of contact days).

[i] Westat (2018). Peer Group Connection-High School: A Cross-Age Peer Mentoring & High School Transition Program. Austin, TX: Westat.

[ii] Grade Point Average (GPA) scales are not consistent across schools and districts. To address this, Westat transformed all GPA scales into a rank order within each district. All students were pooled and sorted on their district GPA from greatest to least. Each student was then assigned their resulting percentile rank within the distribution.

[iii] Impacts for OTI could not be estimated for North Carolina students due to insufficient comparison data.

[iv] Impacts for GPA could not be estimated for North Carolina students due to insufficient comparison data.

CENTER FOR SUPPORTIVE SCHOOLS

AT&T ASPIRE
EVALUATION OF

Peer Group Connection-High School: A Cross-Age Peer Mentoring & High School Transition Program

April 9, 2018

Submitted to

Sherry Barr, PsyD
Managing Director, Operations and Evaluation
Center for Supportive Schools

Submitted by

Westat
An Employee-Owned Research Corporation®
9442 Capital of Texas Highway North
Suite 150
Austin, TX 78759
(301) 251-1500

Contents

Figures

Tables

Executive Summary

In 2014, AT&T contributed nearly $12.5 million in an initiative to support meeting GradNation's goal of a 90% high school graduation rate by 2020. Through the AT&T Aspire initiative, 30 projects designed to serve high school students across the country were funded. In an effort to understand the value of AT&T's investment and the impact these programs have on student outcomes, AT&T contracted with Westat Inc., an independent evaluator. Westat developed a comprehensive data-driven approach to estimate the effect of the 30 AT&T Aspire-funded programs on high school student outcomes, over two academic years (2014–15 and 2015–16). Westat, a large employee-owned, global research firm based out of Rockville, MD, conducted an impact analysis over two academic years (2014–15 and 2015–16) of the 30 AT&T Aspire-funded projects. Students who participated in AT&T Aspire-funded programs were compared to nonparticipating students on the following student outcomes: academic performance, defined as on track to graduate on time (OTI)[1] and grade point average (GPA), and school engagement, defined as school attendance. Additionally, Westat evaluated student program participation, demographic characteristics, participation intensity, and social return on investment (SROI).

Through a highly competitive proposal process, AT&T selected the Center for Supportive Schools' (CSS) Peer Group Connection-High School (PGC-HS) cross-age peer mentoring and high school transition program as two (2) of the 30 projects chosen for participation in the AT&T Aspire initiative. AT&T funded the implementation of PGC-HS in two separate projects: (1) two urban school districts, Baltimore City Public Schools (BCPS) in Maryland and the New York City Department of Education (NYCDOE) in New York, and (2) a rural school district, Sampson County Public Schools, in North Carolina. AT&T invested over one million US Dollars (US$ 1,152,192) combined for the two CSS projects, over two academic years (Year 1: 2014–15 and Year 2: 2015–16). Throughout this report, the project that operated PGC-HS in Baltimore and New York City is referred to as the Baltimore/NYC Project, and the project that operated PGC-HS in the Sampson County Public Schools is referred to as the North Carolina Project.

The *Baltimore/NYC Project* was composed of a total of 28 high schools (9 in Baltimore; 19 in New York City) that implemented PGC-HS during the two years of the AT&T Aspire-funded project. The *North Carolina Project* was composed of four high schools in Sampson County Public Schools that implemented PGC-HS during the two-year grant period.

PGC-HS is an evidence-based, cross-age peer mentoring program that supports and eases the critical period of students' transition from middle to high school. PGC-HS was developed by the Center for Supportive Schools (CSS) (http://www.supportiveschools.org), a nonprofit organization that works in partnership with schools. Since 1979, CSS has partnered with over 200 high schools to implement PGC-HS. PGC-HS taps into the power of older students to create a nurturing environment for incoming freshmen. PGC's mechanism for change is to train select school faculty to prepare older high school juniors and seniors to become peer leaders who mentor and educate freshmen.

This report summarizes the individual program level evaluation results for Baltimore/NYC, and North Carolina, out of the larger AT&T Aspire evaluation. Results for both projects are based on a subsample of students served by the Baltimore/NYC and North Carolina projects with sufficient data provided by affiliated school districts to be included in the matched-comparison analytic sample.

[1] The on-track indicator (OTI) components for being "on track to graduate on time" are (1) core academic course failures and (2) number of course credits to be promoted to the next grade level. The number of course credits required for promotion varied by district, and by grade.

Total Sample of Students Served

The Baltimore/NYC Project and North Carolina Project served 8,179 students over two academic years of programming (Year 1: 2014–15 and Year 2: 2015–16).

- Baltimore/NYC served 3,407 students in Year 1 and 3,290 students in Year 2, for a total of 6,697 students over two years.
- North Carolina served 755 students in Year 1 and 727 students in Year 2, for a total of 1,482 students over two years. A small number of students (*n*=23) participated in both years of programming.

Analytic Samples

In order to identify appropriate comparison groups and conduct the impact analyses, student level data were collected from districts affiliated with the CSS projects. Sufficient data was provided for 5,070 of the 8,179 students served over two academic years of programming (Year 1: 2014–15 and Year 2: 2015–16 academic years) for the Baltimore/NYC Project and the North Carolina Project. Approximately 38% of students served could not be included in the analytic sample due to insufficient data provided by districts to link data and to conduct matching. However, 62% of the students served could be included in the analytic sample.

- Baltimore/NYC: 1,802 students from Year 1 and 1,491 students from Year 2 are included in analytic sample.
- North Carolina: 746 students from Year 1 and 704 students from Year 2 are included in analytic sample.[2]

Research Questions

Five research questions (RQ) were evaluated:

RQ1 What is the impact of participation in the PGC-HS program on on-track to graduation (OTI) status as compared to peer-matched non-program students?

RQ2 What is the impact of participation in the PGC-HS program on grade-point average as compared to peer-matched non-program students?

RQ3 What is the impact of participation in the PGC-HS program on school attendance as compared to peer-matched non-program students?

RQ4 Do students who participate in the top 50% of PGC-HS program intensity experience better outcomes than those participating in the lower 50% of PGC-HS program intensity?

RQ5 What is the lifetime Social Return on Investment (SROI) for each program based on the number of students with better OTI outcomes than their matched controls?

Impact Results

In order to estimate the effect of the PGC-HS program on student outcomes (OTI, GPA, and school attendance), Westat matched program participants with a comparison group of students who did not participate in the program. This matching helped establish equivalent groups prior to program participation, meaning the students in the program were similar/equivalent to their peers in the comparison group before the program began. Any differences in student outcomes following participation in PGC-HS can be attributed to participation in the program, rather than differences in other student characteristics. The following impact results are based on the analytic samples for each project:

- The Baltimore/NYC Project showed statistically significant and positive student outcome results including:
  - ↗ **Increased school attendance** for Year 1 and Year 2 program participants (PGC-HS participants attended school 6.1 more days in Year 1 and 4.7 more days in Year 2 than non-participants)
  - ↗ **A sustained impact on school attendance** for Year 1 program participants was observed at the conclusion of Year 2, a full-year after leaving the program (PGC-HS participants attended school 7.8 more days in Year 2 than non-participants)
  - ↗ **Large effect size** (1.31) for the on track to graduate on time (OTI) indicator for Grade 9 students in Year 1
  - ↗ **Substantively important effect size** (.26) for the on track to graduate on time (OTI) indicator for Grade 10–12 students in Year 1 and a moderate effect size (.52) in Year 2

[2] A small number of students (*n*=23) participated in both years of programming, but the sample was too small to estimate impacts.

- ↗ **A sustained impact on OTI** (effect size = .59) for Grade 9 students was observed for Year 1 program participants at the conclusion of Year 2, a full-year after leaving the program
- ↗ **Higher mean grade point average (GPA)** rank (1.3%) for Year 2 program participants

- The North Carolina Project showed statistically significant and positive student outcome results including:
  - ↗ **Increased school attendance** for Year 1 program participants (PGC-HS participants attended school 6.2 more days than non-participants)
  - ↗ **Increased school attendance** for Year 2 program participants (PGC-HS participants attended school 2.3 more days than non-participants)
  - ↗ **A sustained impact on school attendance** for Year 1 program participants was observed at the conclusion of Year 2, a full-year after leaving the program (PGC-HS participants attended school 2.4 more days than non-participants)

Other outcomes in North Carolina could not be evaluated due to insufficient data.

Participation Intensity Results

The level of student contact with a program can be a critical component of program effectiveness; therefore AT&T Aspire programs were required to report program participation every month for every student. For the purpose of using a common metric across all programs in the AT&T Aspire evaluation, the participation unit was defined as the *day*. Every day that a student was offered (or received) any type of program support or interaction with the program, if even for only 5 minutes, was considered a *participation day*.

Common sense suggests that the more times a student is exposed to effective program activities, the greater the likelihood of positive student impacts, such as increased OTI, GPA, and school attendance. Westat examined differences in intensity within the program. Students were divided into two groups within each program based on the median number of days of participation. Students with participation day totals higher than the median of all students are considered *high-intensity* students and those at or below the median were considered *low-intensity* students. Program impacts were compared between the *high-intensity* and *low-intensity* subgroups.

- In Baltimore/NYC, for Year 1, students participating in the *high-intensity* subgroup (top 50% of contact days) had statistically significantly **higher outcomes in OTI, GPA, and school attendance** than students in the *low-intensity* subgroup (lower 50% of contact days).
- In North Carolina, grade 9 students participating in the *high-intensity* subgroup (top 50% of contact days) had statistically significantly **higher outcomes for OTI** than students in the *low-intensity* subgroup (lower 50% of contact days).

Social Return on Investment (SROI)

As the AT&T Aspire community succeeds, the return on AT&T's investment increases and students benefit by improving attendance in school, improving academics, staying on track in school, graduating on time, and, ultimately, contributing to the welfare of their community by improving their earning potential. Westat estimated the net number of students with favorable OTI outcomes and conducted a SROI analysis to estimate the conservative and optimistic lifelong benefit these students would have on society.

- For Baltimore/NYC, based on the estimated net 356 students with a favorable OTI outcome, the **estimated lifetime SROI is between 11,423% and 36,728%, with a total lifetime social benefit estimated between $111 million and $356 million.**

Discussion

The AT&T Aspire award to CSS represented a unique opportunity, as funding included an external evaluation of the program. The findings reported above demonstrate that the program is working to improve short-term annual outcomes (OTI, GPA and attendance) that are represented in both the PGC-HS and AT&T Aspire evaluation logic models. Therefore, the AT&T Aspire evaluation provides evidence that CSS has successfully implemented the program during the 2014-15 and 2015-16 academic years. This suggests that longer term outcomes (such as on-time four-year graduation and post-secondary success) might be observed as well.

Particularly impressive is the finding for both programs; that Year 1 students see *sustained advantages for school attendance in Year 2* even when no longer attending PGC-HS.

North Carolina and the New York/Baltimore areas represent two very different instructional contexts.

The fact that CSS successfully demonstrated impacts in both regions is impressive. This finding suggests that PGC-HS might scale successfully to other contexts as well; a hypothesis which would need to be validated through future research.

This evaluation found that program intensity, as measured by the number of days of student contact, made a difference in both regions. Students with more program contact time had higher outcomes. CSS may consider working with districts to schedule more program-student time, and possibly, design a study to identify the appropriate number of student contact hours by systematically varying the program intensity across schools.

This evaluation did not distinguish between mentors and mentees when examining outcomes. Also, quasi-experimental (QED) matched evaluations are not as rigorous as a randomized study can be, because QED studies can only control for observed variables. Nevertheless, this evaluation observed positive outcomes in both regions, and for both academic years. This type of replicated finding provides stronger support that PGC-HS is working as intended to result in on-time graduation from high school.

AT&T Aspire Evaluation of Peer Group Connection-High School: A Cross-Age Peer Mentoring & High School Transition Program

Background Information

In 2014, AT&T contributed nearly $12.5 million in an initiative to support meeting GradNation's goal of a 90% high school graduation rate by 2020. Through the AT&T Aspire initiative, 30 projects designed to serve high school students across the country were funded. In an effort to understand the value of AT&T's investment and the impact these programs have on student outcomes, AT&T contracted with Westat Inc., an independent evaluator. Westat developed a comprehensive data-driven approach to estimate the effect of the 30 AT&T Aspire program on high school student outcomes, over two academic years (2014–15 and 2015–16). Westat, a large employee-owned, global research firm based out of Rockville, MD, conducted an impact analysis of 30 AT&T Aspire projects comparing AT&T Aspire students to nonparticipating students on students' academic performance, defined as on track to graduate on time (OTI)[3] and grade point average (GPA), and school engagement, defined as school attendance. Additionally, Westat evaluated student program participation, demographic characteristics, participation intensity, and social return on investment (SROI).

Through a highly competitive proposal process, AT&T selected the Center for Supportive Schools' (CSS) Peer Group Connection-High School (PGC-HS) program as two (2) of the 30 projects chosen for participation in the AT&T Aspire initiative.

AT&T funded the implementation of PGC-HS in two separate projects: (1) two urban school districts, Baltimore City Public Schools and the New York City Department of Education (NYC), and (2) a rural school district in North Carolina, Sampson County Public Schools. The programs were funded for two years (Year 1: 2014–15 and Year 2: 2015–16 academic years). AT&T invested over one million US Dollars (US$ 1,152,192) combined for the two CSS programs, over two-years (Year 1: 2014–15 and Year 2: 2015–16 academic years). Throughout this report, the New York City/Baltimore CSS location is referred to as the Baltimore/NYC Project, and the Sampson County location is referred to as the North Carolina Project.

The Baltimore/NYC Project was composed of a total of 28 high schools (9 in Baltimore; 19 in New York City) that implemented PGC-HS during the two years of the grant-funded project. The *North Carolina Project* was composed of four high schools in Sampson County Public Schools that implemented PGC-HS during the two-year grant period.

Peer Group Connection-High School (PGC-HS) is an evidence-based, cross-age peer mentoring program that supports and eases the critical period of students' transition from middle to high school. PGC-HS was developed by the Center for Supportive Schools (CSS) (http:// www.supportiveschools.org), a nonprofit organization that works in partnership with schools. Since 1979, CSS has partnered with over 200 high schools to implement PGC-HS.

[3] The on-track indicator (OTI) components for being "on track to graduate on time" are (1) core academic course failures and (2) number of course credits to be promoted to the next grade level.

This report summarizes the individual program level evaluation results for the Baltimore/NYC Project, and the North Carolina Project, out of the larger AT&T Aspire evaluation. Results for both programs are based on a subsample of students served by the Baltimore/NYC and North Carolina Projects with sufficient data provided by affiliated school districts, to be included in the matched-comparison analytic sample.

AT&T Aspire Evaluation Logic Model, Evaluation Framework, and Common Outcomes

The AT&T Aspire program seeks to fund high-school completion programs across the country for the purpose of achieving Grad Nation's goal of a 90% on-time high school graduation rate by 2020.[i]

AT&T believes in funding evidence-based programs. Evidence was a requirement during the RFP process and evidence was generated during the funding period. A benefit of the AT&T Aspire program is that external evaluation is provided in addition to the program budget.

Given the variety of student support services offered across all AT&T Aspire programs, AT&T and Westat developed a *common set of outcomes* for all programs that met the following criteria: (1) they would be available for all students in a public school, (2) they would likely be sensitive over one academic year to programmatic activities of programs, and (3) they represent near-term (i.e. proximal) outcomes that lead to the ultimate goal, on-time graduation.

AT&T and Westat selected *on-track to on-time graduation* (OTI), *grade-point average* (GPA), and *school attendance* (Attendance) as the common set of outcomes across programs. In addition, prior AT&T Aspire evaluations have shown program intensity as an important mediator of these outcomes resulting in better annual outcomes; therefore, program participation was collected on a monthly basis. Further details about each of these outcomes are described in the sections below.

Research Questions

Five research questions (RQ) were asked in the AT&T Aspire evaluation:

RQ1 What is the impact of participation in the PGC-HS program on on-track to graduation (OTI) status as compared to peer-matched non-program students?

RQ2 What is the impact of participation in the PGC-HS program on grade-point average as compared to peer-matched non-program students?

RQ3 What is the impact of participation in the PGC-HS program on school attendance as compared to peer-matched non-program students?

RQ4 Do students who participate in the top 50% of PGC-HS program intensity experience better outcomes than those participating in the lower 50% of PGC-HS program intensity?

RQ5 What is the lifetime Social Return on Investment (SROI) for each program based on the number of students with better OTI outcomes than their matched controls?

On-track to On-time Graduation (OTI)

OTI is one of the two academic outcome measures used in the AT&T Aspire evaluation. This outcome was adopted from the University of Chicago's Consortium of School Research, developed for Chicago Public Schools.[ii] The measure was developed specifically for use at the end of 9th grade, to predict whether a student was on-track for on-time graduation. As shown in figure 1 below, the way the OTI metric is constructed is directly related to grade progression; therefore, the OTI was used as an outcome for every grade 9-12. In this way, students from multiple grades could be combined in a single analysis.

The OTI is constructed as follows:

First, the number of credits accumulated during the academic year is calculated. Second, the number of semester-level failures in *core courses* is counted. These two variables taken together (see Figure 1) determine whether a student is considered to be *on-track* or *off-track* for on-time graduation in four years. Note that 12 semester-credits are required in this example district to move from one grade to another. This often varies by district and grade, therefore the actual number of credits required to move from the current grade to the next was acquired from each district and used in the calculation.

Figure 1 | On-track to Graduation Indicator Matrix

	2+ Core Failures	0-1 Core Failures
<12 Credits	Off-track	Off-track
12+ Credits	Off-track	On-track

Grade Point Average (GPA)

Grade Point Average (GPA) is the second of the two academic outcome measures used in the AT&T Aspire evaluation. GPA scales are not consistent across schools or districts, and sometimes can vary greatly even within the *same* district and school. For example, some districts use a 0–99 scale, others an A–F scale (0.0–4.0), and still others add additional GPA points for Advanced Placement courses resulting in a 0.0–4.5 or 0.0–5.0 scale. Also challenging is that middle school GPAs often mean something different than they do in high school. Therefore, 8th grade baseline data for 9th grade students may not necessarily represent the same thing. In order to overcome these variations in GPA scales, Westat transformed all the GPA scales into a rank order within each district. All program and non-program students were pooled and sorted on their district GPA from greatest to least. Each student was then assigned their resulting percentile rank in this distribution.

School Attendance (Attendance)

School attendance is often used by researchers as a generic measure of *student engagement*. All schools must report it and it is a convenient measure to compare across districts. While states, districts, and schools may offer/require a different number of days, actual attendance is easily transformed into a percentage of attendance (number of days actually attended divided by the number of school days offered). This attendance percentage is the common metric used.

Program Participation Intensity

The level of student contact with a program can be a critical component of program effectiveness, as evidenced in prior AT&T Aspire research. Therefore, AT&T Aspire programs were required to report program participation every month for every student.

For the purpose of using a common metric across all programs in the AT&T Aspire evaluation, the participation unit was defined as the *day*. Every day that a student was offered (or received) any type of program support or interaction with the program, if even for only 5 minutes, was considered a *participation day*.

To serve as a proxy for program intensity, two metrics were collected for each student each month: the number of days in that month that a program support was *offered* (planned or simply open and available to the student), and the number of days during that month that a student *actually received* services. A third metric (*participation rate*) was calculated from these by dividing the number of days *received* by the number of days *offered*. Each program entered this data into an Excel template and uploaded it to a secure data sharing folder every month.

Budget Information – Social Return on Investment

Another outcome collected as part of the AT&T Aspire evaluation are actual budget expenditures. The total amount of AT&T Aspire funds expended annually is divided by the number of students represented in the OTI outcome. Standard economic assumptions are made about individual life-time earnings between a non-graduate and a graduate of high-school. These assumptions are multiplied by the number of positive OTI outcomes to estimate a lifetime Social Return on Investment (SROI) resulting from AT&T's investment in the program.

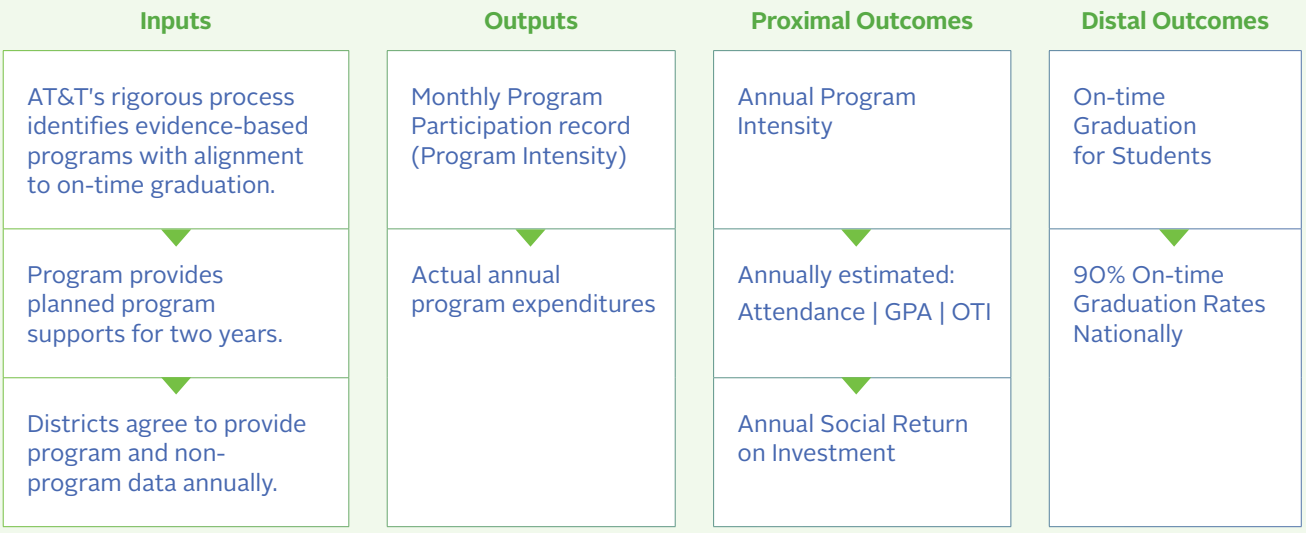
AT&T Aspire Evaluation Logic Model

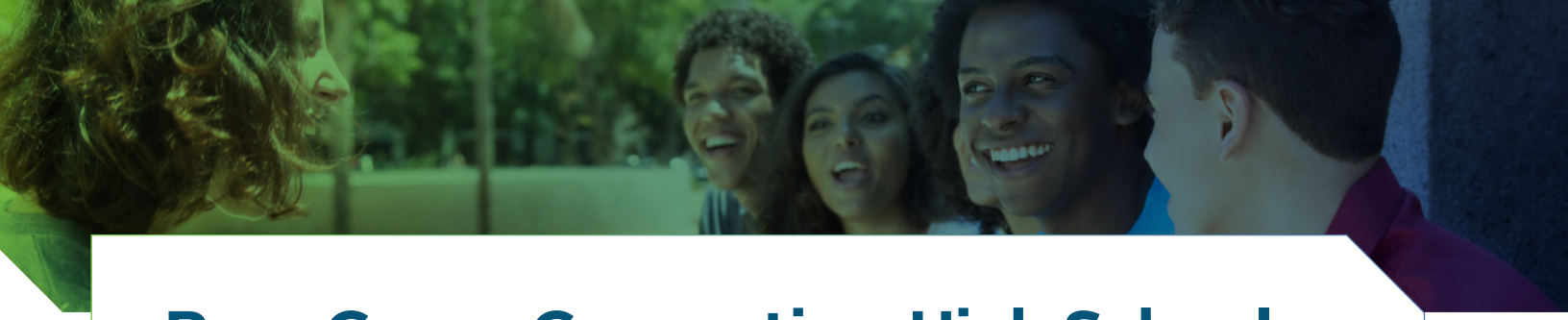
The AT&T Aspire evaluation logic model, as depicted in Figure 2, shows the relationship between the primary output of each program (student participation in programmatic activities) and the near-term outcomes (OTI, GPA, and Attendance). These outcomes are considered proximal because it is thought that they can be changed within a single academic year. They are the primary outcomes of the evaluation because AT&T Aspire programs are funded for two-years, not four. Programs serving 9th and 10th graders would not be able to observe outcomes related to actual four-year on-time graduation.

The most important idea in the AT&T Aspire evaluation's theory of action is that the relationship between programmatic activities and proximal outcomes is mediated by **intensity**. Program intensity has been operationalized (i.e. measured) by monitoring the number of days a program has had any contact with a student.

The next section of the report describes the structure, theoretical framework, and logic model of the Center for Supportive Schools' (CSS) PGC-HS program.

Figure 2 | AT&T Aspire Evaluation Logic Model





Peer Group Connection-High School (PGC-HS): A Cross-Age Peer Mentoring and High School Transition Program

Peer Group Connection-High School (PGC-HS) is an evidence-based, cross-age peer mentoring program that supports and eases the critical period of students' transition from middle to high school. PGC-HS was developed by the Center for Supportive Schools (CSS) (http:// www.supportiveschools.org), a nonprofit organization that works in partnership with schools. Since 1979, CSS has partnered with over 200 high schools to implement PGC-HS.

Purpose and Scope

There is a profound weakness in the support provided to students during the transition into high school. By the time they reach high school, as many as 40 to 60 percent of all students—urban, suburban, and rural— are "chronically disengaged" from school.[iii] Disengagement is often associated with behavior and discipline problems and may eventually lead to dropout.[iv] Nearly 70 percent of the high school dropouts interviewed in the landmark 2006 study *The Silent Epidemic: Perspectives of High School Dropouts* said that not feeling motivated or inspired to work hard was a major factor in their decision to drop out.[v]

Research also consistently demonstrates that students are most vulnerable for dropping out of school during and immediately following their first year of high school.[vi] More students fail ninth grade than any other grade and promotion rates between ninth and tenth grades are much lower than rates between other grades.[vii,viii] By focusing on the transition from middle to high school, this period of heightened vulnerability has the potential for being transformed into a window of opportunity for preventing the consequences of disengagement and poor decisions.

PGC-HS taps into the power of older students to create a nurturing environment for incoming freshmen. PGC-HS utilizes peer leaders to increase school engagement, develop social and emotional learning skills, and reinforce desirable behaviors resulting in improved educational outcomes, including improved graduation rates.

Program Structure

PGC-HS is implemented in high schools (grades 9–12). PGC-HS's **mechanism for delivery** is to train select school faculty to prepare high school juniors/seniors to mentor and educate freshmen and create a positive school environment:

1. PGC-HS's launch begins with the assembly of a ***stakeholder team*** of administrators, the school scheduler, and other faculty and staff, and is led by a coordinator, who receive the training, tools, and resources necessary to meet regularly to plan for implementation of PGC-HS, troubleshoot obstacles, and ensure PGC-HS's long-term sustainability.

2. Faculty members are carefully selected by the stakeholder team to serve as ***faculty advisors***. CSS provides the stakeholder team with tools and guidance to select faculty advisors, including assessing for criteria such as: enthusiasm for PGC-HS and a peer leadership approach; ability to work collaboratively with others; willingness to implement PGC-HS with fidelity; and willingness to utilize a facilitation model rather than didactic teaching style. Faculty advisors participate in an 11-day intensive train-the-trainer course over a 1½-year period to learn how to run the program and teach junior and senior peer leaders in a daily leadership class (see #4).

3 Juniors and/or seniors are carefully selected by faculty advisors to become ***peer leaders*** for 9th graders. CSS supports faculty advisors to select a diverse group of peer leaders that reflects the racial/ethnic composition of the school community, neighborhood affiliation, socio-economic status, known cliques, and an equal number of girls and boys. CSS provides faculty advisors with tools and guidance to recruit and select peer leaders, including recruitment tools that make the opportunity to apply to serve as a peer leader available to all students in the targeted grade. Selection is based on applications, faculty recommendations, interviews, school performance, and criteria such as clear commitment to the role of peer mentor, ability to work well in groups, enthusiasm, and supportive of others. Standard protocols and training for each of these components are provided to all PGC-HS schools.

4 Peer leaders are trained and conduct weekly outreach sessions as part of their regular school schedule in a ***daily, 45-minute leadership development class*** typically offered as an elective course for credit. Within the daily class, peer leaders receive 4 days of training for every 1 day of mentoring they provide to freshmen. This helps peer leaders prepare to lead their groups each week and debrief following each session, sharing successes and challenges, and problem-solving.

5 Peer leaders work in pairs to co-lead ***peer groups***, small groups of 10-14 freshmen. PGC-HS typically replaces one day per week of physical education for freshmen in participating high schools. CSS also provides a structured, 2-hour protocol for co-leader selection and assignment to lead peer groups. CSS works closely with administrators and staff beginning prior to program implementation to coordinate scheduling.

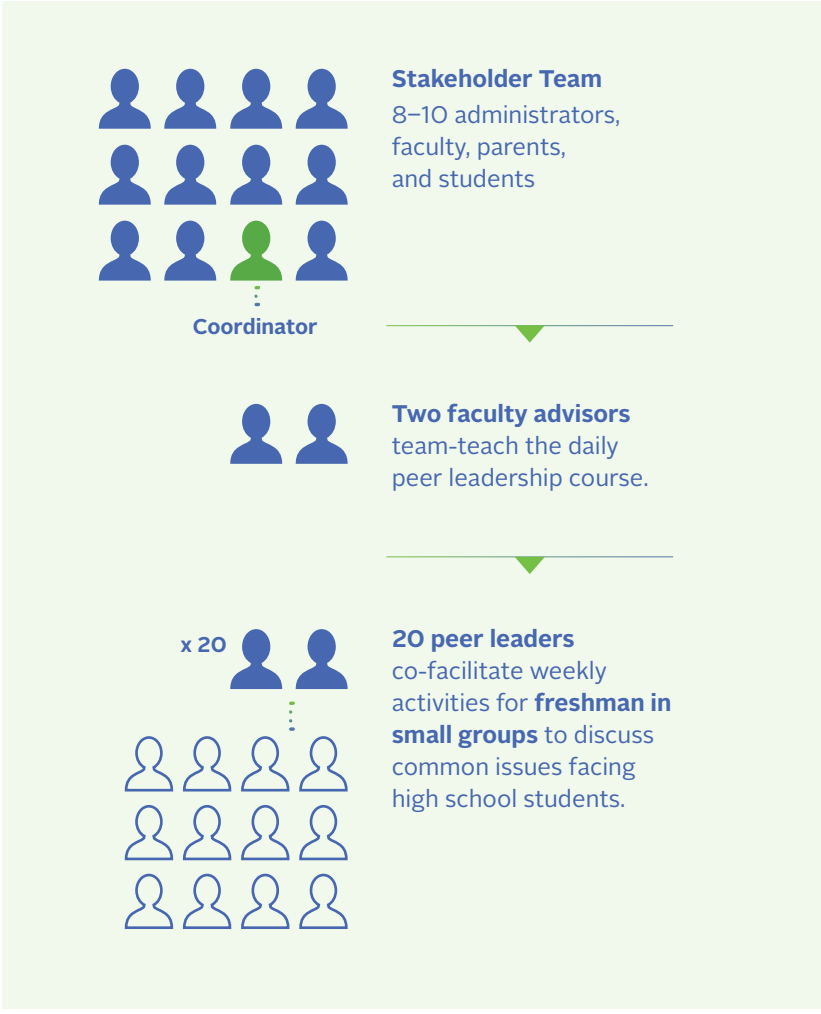
6 Peer leaders work in pairs to co-lead groups of 10 to 14 freshmen in ***weekly outreach sessions*** during the school day in which the freshmen participate in engaging, hands-on activities and discussions on a variety of youth development topics such as sense of school attachment, competence in interpersonal relationships, conflict resolution, motivation, goal setting, coping skills, decision making, growth mindset, peer acceptance, and resisting peer pressure.

7 Peer groups research, plan, and execute a ***service learning project***, using a structured framework to support meaningful, youth-led community involvement through a multi-layered action research model.

8 PGC also includes a ***family involvement*** component. Peer leaders organize and facilitate a Family Night event for freshmen and their families that is focused on increasing family-teen communication and showcasing community service projects.

The PGC-HS program structure is depicted in Figure 3, below.

Figure 3 | Peer Group Connection – High School Program Structure



Theoretical Framework

PGC-HS is grounded in social and emotional learning (SEL) theories, which can play a valuable role combatting young people's negative perceptions. The mindsets model of SEL suggests that environments can socialize children and adolescents to hold different belief systems, or mindsets.[ix] These mindsets in turn cause them to use (or not use) the skills that they have or are acquiring. When SEL programs offer adolescents a route to feelings of status and respect, it is more likely that they will internalize acquired skills and apply them in the real world. SEL programs like PGC-HS try to help adolescents cope with their difficulties more successfully by improving skills and mindsets while simultaneously creating respectful school environments.[x] Evidence-based SEL programs can enhance children's confidence in themselves, increase their engagement in school, and promote desirable behaviors.[xi] Further, SEL can make a profound difference for students in disadvantaged communities because it has the potential to create a safe classroom environment in which students and educators can have open, honest, and validating conversations about the reality of what students face every day.[xii] It also can provide students with emotional tools to counter negative messages and stand up against racism and marginalization in their communities.[xiii] SEL can help students from disadvantaged backgrounds overcome the cognitive and emotional scars they suffer as a result of their environment by highlighting the strengths and challenges that an individual or community are bringing to the table.[xiv] A mounting body of evidence clearly indicates that, compared to students who do not participate in such programs, students who receive SEL programming academically outperform their peers, get better grades, and graduate at higher rates.[xv] SEL core competencies are fully addressed within the PGC curriculum.

Social network[xvi] and social learning[xvii] theories are also important. Social network theory suggests peer relational structures play a role in students' access to social and academic resources and information.[xviii,xix] Theoretically, students in interconnected and heterogeneous peer groups, with central members who are academically-oriented and prosocial, have greater access to the information and support necessary to succeed in school.[xx] PGC-HS involves diverse students from different levels of risk for school dropout in same-age peer groups with older peer leaders. Lower-risk students, who demonstrate fewer overt signs of distress, receive same-age and cross-age support to overcome obstacles that could eventually lead to more serious problems. Youth at both moderate and high risk for dropout benefit from exposure to more academically-oriented students in a supportive, structured setting and from observational learning and imitation of these peer models.[xxi,xxii]

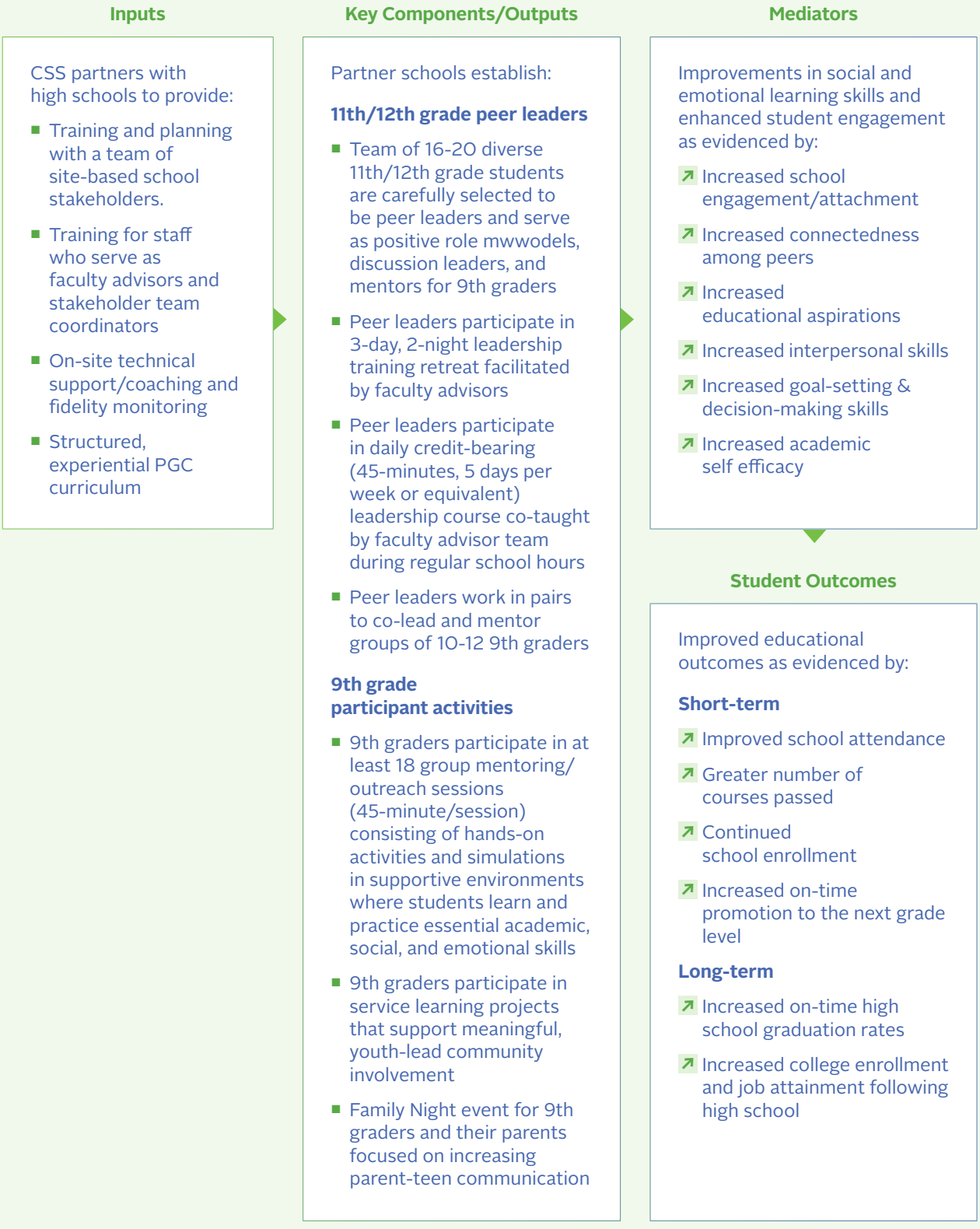
Logic Model

PGC-HS's logic model is depicted in Figure 4.

Alignment of the PCG-HS and AT&T Aspire Evaluation Logic Models

A comparison of the AT&T Aspire evaluation model and the PCG-HS logic model reveals a very close alignment between the two. Note that two of the AT&T Aspire proximal outcomes (OTI & Attendance) are identical to PGC-HS short term outcomes. AT&T Aspire's distal outcome (on-time high-school graduation) is identical to the first PGC-HS long-term outcome. AT&T Aspire's *participation intensity mediator* is represented in terms of a precisely stated participation dosage in the key components/outputs section of the PGC-HS logic model. This alignment suggests that the findings of the AT&T Aspire evaluation are directly relevant to an understanding of how effective PCG-HS is in achieving its intended outcomes.

Figure 4 | Peer Group Connection – High School Logic Model



Results of the AT&T Aspire Evaluation of the PGC-HS Program

The AT&T Aspire evaluation is focused on three outcomes: the *on track to graduation indicator* (OTI) and *GPA* (both representing academic performance), and *school attendance* (Attendance, representing school engagement). The evaluation took place over two years (Year 1: 2014–15 and Year 2: 2015–16 academic years).

CSS was awarded funding for their Peer Group Connection-High School (PGC-HS) program in two distinct community settings: (1) two urban school districts, Baltimore City Public Schools (BCPS) in Maryland and the New York City Department of Education (NYCDOE) in New York, and (2) a rural school district in North Carolina, Sampson County Public Schools. Throughout this report, the urban school districts are referred to as the *Baltimore/NYC Project*, and the rural school district is referred to as the *North Carolina Project*.

The *Baltimore/NYC Project* was composed of a total of 29 high schools (9 in Baltimore; 19 in New York City) that implemented PGC-HS during the two years of the grant-funded project. The *North Carolina Project* was composed of four high schools in Sampson County Public Schools that implemented PGC-HS during the two-year grant period.

Because the PGC-HS program participants primarily participate for one year, not both years of the two-year AT&T grant period,[4] this resulted in three analytic groups as described next.

Year 1 only students at the end of year 1. This group includes all students in the PGC-HS program in Year 1 (2014–15 academic year), with impacts estimated at the end of the academic year.

Year 2 only students at the end of year 2. This group includes all students who were new to the PGC-HS program in Year 2 (2015–16 academic year), with impacts estimated at the end of the academic year.

Year 1 only students at the end of year 2. This group includes all students in the PGC-HS program in Year 1 (2014–15 academic year) and received only one year of support. These are impacts estimated at the end of Year 2 after a delay or lag of one year following the end of program participation.

Only students who could be linked with their district data were included in the analysis. The demographic characteristics of Year 1 and Year 2 students are provided in Table 1 for both the Baltimore/NYC and North Carolina projects. The demographic characteristics of Year 1 students measured at the end of Year 2 are provided for both projects in Table 2. The first row of each table provides the total count of students participating in the program. The second row of each table provides the count and proportion of student subgroups for whom demographic and other data could be linked from a district.

[4] N=23 students participated in the program in North Carolina for both years (i.e. Year 1 and 2 students). This was not enough to estimate stand- alone impacts.

Table 1 | Demographics for the Students With Available District Data, Year 1 and Year 2

		Year 1 Only				Year 2 Only			
		NYC/Baltimore		North Carolina		NYC/Baltimore		North Carolina	
Total N From Participation Counts		3407		755		3290		727	
N Count and % With District Link		1802	52.9%	746	98.8%	259	7.8%	704	96.8%
Demographic Variable	Value	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent
Grade	Missing	0	0	61	8	3	1	54	0
	9	1403	78	604	81	7	3	599	85
	10	68	4	0	0	7	3	4	1
	11	156	9	48	6	85	32	37	5
	12	175	10	33	4	160	61	10	1
Gender	Missing	0	0	11	1	3	1	35	5
	Male	834	46	381	51	96	37	347	49
	Female	968	54	354	47	163	62	322	46
Race/Ethnicity	Missing	6	0	3	0	5	2	45	6
	American Indian	13	1	8	1	2	1	3	0
	Asian/Pacific Islander	137	8	5	1	11	4	1	0
	African American	523	29	160	21	116	44	155	22
	Hispanic	1060	59	216	29	116	44	191	27
	White	58	3	336	45	12	5	293	42
	Two or More Races	5	0	18	2	0	0	16	2
Limited English Proficient	Missing	0	0	746	100	3	1	174	25
	No	1503	83	0	0	244	93	507	72
	Yes	299	17	0	0	15	6	23	3
Special Education	Missing	0	0	746	100	3	1	175	25
	No	1431	79	0	0	229	87	487	69
	Yes	371	21	0	0	30	11	42	6
Economically Disadvantaged	Missing	0	0	746	100	3	1	704	100
	No	288	16	0	0	39	15	0	0
	Yes	1514	84	0	0	220	84	0	0

Table 2 | Demographics for the Students With Available District Data, Year 1 Students Measured at the End of Year 2

		Year 1 Sustained			
		NYC/Baltimore		North Carolina	
Total N From Participation Counts		3407		755	
N Count and % With District Link		1818	53.4%	732	97.0%
Demographic Variable	Value	Counts	Percent	Counts	Percent
Grade	Missing	267	15	75	0
	9	209	12	71	10
	10	1142	63	544	74
	11	55	3	2	0
	12	145	8	40	5
Gender	Missing	267	15	18	2
	Male	727	40	368	50
	Female	824	45	346	47
Race/Ethnicity	Missing	271	15	19	3
	American Indian	13	1	8	1
	Asian/Pacific Islander	122	7	4	1
	African American	437	24	152	21
	Hispanic	917	50	210	29
	White	54	3	321	44
	Two or More Races	4	0	18	2
Limited English Proficient	Missing	267	15	385	53
	No	1330	73	344	47
	Yes	221	12	3	0
Special Education	Missing	267	15	381	52
	No	1211	67	328	45
	Yes	340	19	23	3
Economically Disadvantaged	Missing	267	15	732	100
	No	294	16	0	0
	Yes	1257	69	0	0

Creating the Analytic Sample

The most rigorous possible evaluation would assign students at random to participate in PGC-HS or not to participate in PGC-HS. Random assignment can reduce the possibility that unmeasured variables might bias the impact results. However, with AT&T Aspire funded programs, random assignment was not possible.

A matched-comparison design was, therefore, used to organize the data and produce an estimate of the effect of program participation on participants' OTI, GPA, and school attendance. There will always be reservations about potential bias in outcomes estimates because unmeasured pre-program differences could not be controlled for and may be included in the estimate.

A propensity score, estimated through logistic regression, combines information from many baseline variables into a single value. That single value is estimated using both program and non-program students combined. Once estimated, the propensity score is used to match each program student to one or more non-program students.

We attempted to estimate different outcomes for each of the student groups. For OTI, grade 9 students are estimated separately from grades 10–12. This is because grade 8 data from middle school (grades and credits), which are used for matching, are often less trustworthy and/or missing, which result in high attrition. For *GPA and attendance*, all students in grades 9–12 were combined into one analysis. Students were first matched within grade using PSM. Once the program and non-program matches were identified, matched students across all grades were pooled and used to estimate a single impact estimate.

After matching, baseline equivalence between the matched program and non-program groups was checked. If enough of the baseline covariates were within 0.25 standard deviations (see *Matching Attempt Logic* in Appendix 3), then the groups were considered successfully matched.

Data Acquisition Challenges

Westat prefers to match students to comparison peers within the same district. However, there were substantial challenges with district data acquisition for both projects, which required CSS students

from Baltimore and all of the North Carolina Project students to be matched to peers outside of their districts. Baltimore City Public Schools would only execute a data-sharing agreement to provide data for *program students only*; therefore, Baltimore students were matched to peers from New York City. The North Carolina Project served all 9th grade students in the district; therefore, a suitable age-appropriate comparison group was not available from within the district. North Carolina students were matched to students from a similar district in another Southeastern state.

Attrition Concerns

As the analytic process moves from the participation sample (i.e. the population of PCG-HS students), to the sample of students linked to district demographics, and finally to the impact samples, the available sample sizes become much smaller. In some cases, the impact samples represent fewer than 12% of the original participation sample. This raises the concern of how representative the impact samples are of the total participation sample. While impact estimates are valid for the students included in them, it is important to take caution in generalizing to all students in the program. One must consider what proportion of the original sample is included in the impact estimates and how similar subsamples are to the population both in terms of measured and unmeasured variables. Appendix 1 displays the relationship between the participation samples, the sample of students with district data, and the analytic sample for each year of data in both the Baltimore/NYC and North Carolina Projects.

Impact Findings

All impact findings for both the Baltimore/NYC Project and North Carolina Project are provided in Table 3 (Baltimore/NYC) and Table 6 (North Carolina), below. Demographic summaries of the students represented in each impact are given after each program impact table, in Tables 4–5 (Baltimore/NYC) and Tables 7–8 (North Carolina).

How to Interpret the Impact Tables

Each of the three outcomes (OTI, GPA, Attendance) are expressed in different metrics within the outcome tables, to improve interpretability. All could be reported as effect sizes, which would aid direct comparison, but can obscure the meaning.

OTI is reported as an effect size. An effect size is calculated from the mean and standard deviation of the outcome for the program and non-program groups. Effect sizes of 0.10 and greater are often considered of value. The What Work Clearinghouse considers effect sizes of 0.25 to be of substantive importance even if not significant. Please see Appendix 3 for more explanation of effect size interpretation. For OTI, grade 9 students are estimated separately from grades 10–12.

GPA is expressed in terms of percentile rank. Because GPA scales vary across districts, all GPAs are converted into a within-district percentile rank. A 2.0 difference indicates that the mean GPA of the PGC-HS group is 2% higher than the mean of their matched controls on a 0–100 scale.

Attendance is expressed as a percentage of attended days out of all days offered. An impact of 2.0 equals 2% of the days offered; or for example, if 180 days of school are offered * .02 = 3.6 days mean attendance impact between PGC-HS and control schools.

Baltimore/NYC Impact Project Findings

Impact findings for the Baltimore/NYC Project are provided in Table 3. Of the twelve impacts that can be reported eight are positive and statistically significant. This type of replicable findings in successive years, and especially for students after a year follow-up, is strong evidence that the PGC-HS program is moving the needle in academic outcomes per the program's logic model.

OTI: In Year 1, results among 9th grade students showed a very large effect size difference for OTI, over one standard deviation (1.31), while 10-12th grade students showed a small, but substantively important effect size difference of 0.26. In Year 2, no impacts were noted for OTI among 9th grade students, but a moderate and statistically significant impact (0.52) was observed for 10-12th grade students. Estimates for both years are based on a small subsample of students (Year 1: 11.5% Year 2: 6.1%). Results among Year 1 9th grade program participants assessed at the end of Year 2 showed a sustained moderate effect size difference (.59) for OTI.

GPA: In year 2, Baltimore/NYC students exhibited a 1.3% higher mean GPA rank than matched comparison students. This impact finding was statistically significant.

Table 3 | Baltimore/NYC Project Impacts

Baltimore/NYC	Impact Estimates	Proportion of Participants Represented
Y1 only – End of 2014-15		
OTI 9 (effect size)	↑(1.31)	11.5%[a]
OTI 10-12 (effect size)	↑(0.26)	
GPA (percentage rank of 100)	*NE*	20.9%
Attendance (percentage of days)	↑(3.38)	27.4%
Y2 only – End of 2015-16		
OTI 9 (effect size)	*NE*	6.1%
OTI 10-12 (effect size)	↑(0.52)	
GPA (percentage rank of 100)	↑(1.33)	7.8%
Attendance (percentage of days)	↑(2.61)	7.8%
Y1 Students Persistent Impact in Y2 2015-16		
OTI 9 (effect size)	↑(0.59)	6.1%
OTI 10-12 (effect size)	*NE*	
GPA (percentage rank of 100)	*NE*	5.7%
Attendance (percentage of days)	↑(4.31)	21.0%

[a] The subgroup that could be matched is possibly not representative of all students served.

Key: ↑ indicates a positive and statistically significant impact

NE indicates no statistically significant or substantively important effects

Attendance: Attendance impacts for both Year 1 and Year 2 students was positive and statistically significant. In Year 1, Baltimore/NYC PGC-HS students exhibited a 3.4% higher attendance rate than matched comparison students. Given a 180-day school year, this translates to PGC-HS students attending an average of 6.1 more days than matched comparison students (.0338% x 180 days = 6.08). In Year 2, Baltimore/NYC PGC-HS students exhibited a 2.6% higher attendance rate than matched comparison students, resulting in an average of 4.7 additional attendance days than matched comparison students (.0261% x 180 days = 4.70). A positive and statistically significant impact was also observed for Year 1 PGC-HS students at the end of Year 2, a full year following their participation in the PGC-HS program. Year 1 PGC-HS students assessed at the end of Year 2 exhibited a 4.3% higher attendance rate than matched comparison students, resulting in an average of 7.8 additional attendance days (.0431% x 180 days = 7.76). *This last finding is particularly interesting, as it demonstrates a sustained lasting impact of PGC-HS on students. These students had not participated in the program for an entire year, and still demonstrated increased attendance compared to matched students who did not participate in the PGC-HS program.* All estimates should be interpreted with some caution as they are based on samples ranging from 8% to 27% of all participants and, therefore, may not generalize to the total population of participating students. Attention should be paid to the demographic makeup of these small samples in Tables 4 and 5.

Table 4 | Baltimore/NYC Project Impact Demographics for Year 1 and Year 2

		Year 1								Year 2 Only							
Outcome		Attendance		GPA		OTI9		OTI10-12		Attendance		GPA		OTI9		OTI10-12	
Total N From Participation Counts		3407								3290							
Total N With District Link		1802								1491							
Percent Matched of Participants		932	27.4%	377	11.1%	391		11.5%		256	7.8%	256	7.8%	209		6.1%	
Demographic Variable	Value	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent
Grade	9	579	62	25	7	54	100	0	0	6	2	6	2	8	2	0	0
	10	56	6	55	15	0	0	40	12	7	3	6	2	37	7	1	1
	11	135	14	135	36	0	0	135	40	84	33	84	33	2	0	31	18
	12	162	17	162	43	0	0	162	48	159	62	159	62	0	0	130	73
Gender	Male	456	49	178	47	34	63	154	46	94	37	93	36	31	6	75	42
	Female	476	51	199	53	20	37	183	54	162	63	163	64	16	3	87	49
	Missing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Race/Ethnicity	American Indian	8	1	2	1	3	6	2	1	1	0	2	1	3	1	0	0
	Asian/ Pacific Islander	50	5	25	7	6	11	25	7	11	4	11	4	6	1	13	7
	African American	280	30	120	32	12	22	109	32	116	45	116	45	11	2	41	23
	Hispanic	565	61	220	58	28	52	192	57	116	45	115	45	24	5	105	59
	White	28	3	9	2	5	9	8	2	12	5	12	5	3	1	3	2
	Two or More Races	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
Limited English Proficient	No	804	86	315	84	33	61	285	85	241	94	241	94	32	6	134	76
	Yes	128	14	62	16	21	39	52	15	15	6	15	6	15	3	28	16
Special Education	No	730	78	334	89	54	100	316	94	226	88	226	88	47	9	154	87
	Yes	202	22	43	11	0	0	21	6	30	12	30	12	0	0	8	5
Economically Disadvantaged	No	156	17	71	19	13	24	63	19	37	14	38	15	8	2	27	15
	Yes	776	83	306	81	41	76	274	81	219	86	218	85	39	7	135	76

Table 5 | Baltimore/NYC Project Impact Demographics for Year 1 Persistent Outcome Analyses

		Year 1 Only Sustained							
Outcome		Attendance		GPA		OTI9		OTI10-12	
Total N From Participation Counts		727							
Total N with District Link		3407							
Percent Matched of Participants		1802	52.9%	746	98.8%	259		96.8%	
Demographic Variable	Value	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent
Grade	9	144	20	9	5	8	17	0	0
	10	395	55	10	5	37	79	1	1
	11	44	6	43	22	2	4	31	19
	12	131	18	131	68	0	0	130	80
Gender	Male	357	50	91	47	31	66	75	0
	Female	357	50	102	53	16	34	87	1
	Missing	0	0	0	0	0	0	0	0
Race/Ethnicity	American Indian	7	1	0	0	3	6	0	0
	Asian/Pacific Islander	36	5	13	7	6	13	13	8
	African American	201	28	49	25	11	23	41	25
	Hispanic	448	63	127	66	24	51	105	65
	White	22	3	4	2	3	6	3	2
	Two or More Races	0	0	0	0	0	0	0	0
Limited English Proficient	No	627	88	158	82	32	68	134	83
	Yes	87	12	35	18	15	32	28	17
Special Education	No	541	76	170	88	47	100	154	95
	Yes	173	24	23	12	0	0	8	5
Economically Disadvantaged	No	130	18	30	16	8	17	27	17
	Yes	584	82	163	84	39	83	135	83

Table 6 | North Carolina Project Impacts

North Carolina	Impact Estimates	Proportion of Participants Represented
Y1 Only – End of 2014-15		
OTI 9 (effect size)	X	X
OTI 10-12 (effect size)	X	
GPA (percentage rank of 100)	X	X
Attendance (percentage of days)	↑(3.45)	65.6%
Y2 Only – End of 2015-16		
OTI 9 (effect size)	*NE*	52.2%
OTI 10-12 (effect size)	X	
GPA (percentage rank of 100)	X	X
Attendance (percentage of days)	↑(1.28)	76.6%
Y1 Students Persistent Impact in Y2 2015-16		
OTI 9 (effect size)	X	X
OTI 10-12 (effect size)	X	X
GPA (percentage rank of 100)	X	X
Attendance (percentage of days)	↑(1.34)	57.7%

Key: ↑ indicates a positive and statistically significant impact

NE indicates no statistically significant or substantively important effects

X indicates not enough data provided to estimate impact

North Carolina Project Impact Findings

Impact findings for the North Carolina Project are provided in Table 6, below. Of the four impacts that can be reported, three attendance related outcomes are statistically significant and positive. Impacts for GPA and OTI could not be estimated due to insufficient data and poor matching.

Attendance: Attendance impacts for both Year 1 and Year 2 students was positive and statistically significant. In Year 1, North Carolina PGC-HS students exhibited a 3.5% higher attendance rate than matched comparison students. Given a 180-day school year, this translates to PGC-HS students attending an average of 6.2 more days than matched comparison students (.0345% x 180 days = 6.21). In Year 2, North Carolina PGC-HS students exhibited a 1.3% higher attendance rate than matched comparison students, resulting in an average of 2.3 additional attendance days than matched comparison students (.0128% x 180 days = 2.30). A positive and statistically significant impact was also observed for Year 1 PGC-HS students at the end of Year 2, a full year following their participation in the PGC-HS program. Year 1 PGC-HS students assessed at the end of Year 2 exhibited a 1.3% higher attendance rate than matched comparison students, resulting in an average of 2.4 additional attendance days than matched comparison students (.0134% x 180 days = 2.41). *This last finding is particularly interesting, as it demonstrates a sustained lasting impact of PGC-HS on students. These students had not participated in the program for an entire year, and still demonstrated increased attendance compared to matched students who did not participate in the PGC-HS program.*

This type of replicable findings in successive years, and especially for students after a year follow-up, is strong evidence that the PGC-HS program is moving the needle in terms of school attendance per the program's logic model.

All estimates should be interpreted with some caution as they are based on samples ranging from 57% to 77% of all participants and, therefore, may not generalize to the total population of participating students.

Table 7 | North Carolina Project Impact Sample Demographics

		Year 1		Year 2				Year 1 Only Sustained	
Outcome		Attendance		Attendance		OTI9		Attendance	
Total N From Participation Counts		755		727				755	
Total N With District Link		746	98.8%	704		96.8%		732	97.0%
Percent Matched of Participants		495	65.6%	556	76.6%	401	55.2%	430	57.7
Demographic Variable	Value	Counts	Percent	Counts	Percent	Counts	Percent	Counts	Percent
Grade	9	495	100	554	99	401	100	32	7
	10	0	0	3	1	0	0	397	92
	11	0	0	0	0	0	0	1	0
	12	0	0	0	0	0	0	0	0
Gender	Male	258	53	290	52	223	56	226	53
	Female	229	47	266	48	178	44	203	47
Race/Ethnicity	American Indian	6	1	3	1	0	0	6	1
	Asian/Pacific Islander	2	0	0	0	0	0	2	0
	African American	100	20	134	24	118	29	89	21
	Hispanic	145	29	156	28	120	30	127	30
	White	235	47	246	44	155	39	200	47
	Two or More Races	7	1	14	3	6	2	6	1
Limited English Proficient	No	0	0	403	72	235	59	199	46
	Yes	0	0	18	3	19	5	2	0
	Missing	495	100	136	24	147	37	229	53
Special Education	No	0	0	388	70	226	56	192	45
	Yes	0	0	33	6	28	7	12	3
	Missing	495	100	136	24	147	37	226	53
Economically Disadvantaged	No	0	0	0	0	0	0	0	0
	Missing	0	0	0	0	0	0	0	0
	Yes	495	100	557	100	401	100	430	100

Participation Intensity and Its Relationship to Outcomes

The collection of student-level participation data allows for an examination of potential relationships between intensity (or amount) of program participation and program impacts. Common sense suggests that the more time a student is exposed to effective program activities, the greater the positive student impacts are likely to be.

Table 8 summarizes the number of participation days reported by both the Baltimore/NYC and North Carolina Projects. A primary finding is that both Projects offered a different level of intensity as measured by the number of days participated. As external evaluators, it is unclear whether this truly represents a) differences in the programmatic services actually delivered, b) differences in how each site reported participation, or c) differences in how similar programmatic services were allocated over time. For example, a much smaller number of North Carolina students received services in the spring of each year due to Sampson County Schools' course scheduling structure. In this district, high schools operate on a block schedule (classes meet for approximately 90 minutes with four classes scheduled each day over one semester) rather than a traditional schedule (classes meet for approximately 45 minutes with eight class periods scheduled each day over the full academic year or two semesters). While the total program dosage for participants is designed to be the same whether PGC-HS is implemented within a block or traditional schedule, there are fewer days participated when the program is offered within a block scheduling structure.

Intensity Methods

Intensity analyses were conducted for three outcomes: *OTI, GPA, and Attendance*. The analyses categorized students into *high-intensity* and *low-intensity* groups by splitting the groups at the median participation within each project and examining the different outcomes.

- For the analysis, the analytic sample was used.
- Using the monthly participation files from September to May of each academic year for each student, the total number of raw participation days across the academic year was summed, and the percentage of participation days attended vs. offered was calculated.
- Participation data were merged in the propensity score matching (PSM) file using each student's program ID.
- Using those students for whom participation data could be merged into the PSM file, the *median number of participation days* and the *median participation rate* were determined for each program.

Table 8 | Mean Number of Participation Days by Program by Year

	Year 1				Year 2			
Program	N Count	Days Offered	Days Participated	Participation Rate	N Count	Days Offered	Days Participated	Participation Rate
Baltimore/NYC Project	3407	36.7	25.0	0.68	3290	41.1	22.1	0.54
North Carolina Project	755	20.2	18.9	0.94	727	17.4	16.3	0.94

Note: The North Carolina Project reported a much smaller number of students served in the spring than in the fall of Year 1 and Year 2.

Table 9 | Statistically Significant Impact Estimates for Differences Between High- and Low-Intensity Groups

Program	Outcome	Year	Cohen's d	*p* Value
Baltimore/NYC	GPA 9–12	1	0.59	< 0.000
Baltimore/NYC	Attendance 9–12	1	0.47	< 0.000

Program	OTI Outcome	Year	Odds Ratio	*p* Value
Baltimore/NYC	OTI—9	1	6.2	0.01
Baltimore/NYC	OTI—10–12	1	3.31	0.04
North Carolina	OTI-9	2	2.54	0.001

- Students were considered *high intensity* for *days* if they were at the days median or higher and were considered *high intensity* for *rate* if at the *rate* median or higher. Otherwise students were classified as *low intensity*.
- Within each program, impact estimates were calculated for the *low-intensity group* and the *high-intensity group*, and the difference between them tested.

Median Split on Intensity Within Program

Table 9 summarizes the findings for each statistically significant finding. In Year 1 (2014–15 academic year), intensity makes a difference for all outcomes within the Baltimore/NYC Project. Students in the top 50% of program attendance have statistically higher outcomes than students in the lower 50% of program attendance. Split half analyses were not possible for Baltimore/NYC in Year 2 because the participation data could not be linked to district data.

For OTI grade 9 in North Carolina in Year 2 (2015–16 academic year), students in the top 50% of program attendance did better on the OTI outcome.

The intensity analyses indicate that PGC-HS students in the high-intensity subgroups were associated with higher outcomes in OTI, GPA, and attendance using *raw counts of days*, as compared to students in the low-intensity subgroups. These findings were primary observed for the Baltimore/NYC Project, which may be due to the larger number of mean contact days reported for that program.

We also conducted all high- and low- intensity analyses using the *proportion* of time (a ratio of days attended/days offered) for each project, but no comparison was found to be significant. Therefore, the metric when expressed in *counts of days* attended as compared to using a *ratio of days attended* appeared to pick up some differences in intensity.

Social Return on Investment

The annual program expenditure data, used in conjunction with the results of the OTI evaluation above, provide a rich opportunity to explore the financial context of the AT&T Aspire program.

Social return on investment (SROI) is defined as the net financial benefit of the program (benefit minus costs) divided by the cost of the investment. The benefit from the investment is defined as the social financial benefit to society over the lifetime of the individual who graduates from high school, who otherwise would not. The benefit from investment was estimated based on the forecasted lifetime income of each individual (Appendix 4, Table 18).[xxiii] The benefit ranges from a conservative estimate of $312,800 to an optimistic $1,000,033 per individual.

The social return on investment (SROI) here is calculated using the classic ROI formula (New Education Foundation, 2004):

$$SROI = \frac{Benefit\ from\ investment - Cost\ of\ investment}{Cost\ of\ investment} \times 100\%$$

The first step used to calculate SROI was to estimate each Project's impact on program participants' OTI outcome. The results from this calculation were the basis for all ROI calculations. All impacts were based on counts of program students, compared to counts of comparison students.

A favorable OTI outcome for any AT&T Aspire project can be achieved when:

1. a student has moved from off-track to on-track status or;
2. a student has been prevented from going off track.

$$O_i = [(OTI_T^+ - OTI_C^+) - (OTI_T^- - OTI_C^-)]$$

where O_i is the net number of students with favorable outcomes, which is the number of students with favorable outcomes minus the number of students with unfavorable outcomes; OTI_T^+ is the number of students in the project moving from off track to on track; OTI_C^+ is the number of students in the control group moving from off track to on track; OTI_T^- is the number of students in the program moving from on track to off track; and OTI_C^- is number of students in the control group moving from on track to off track.

- For Baltimore/NYC, based on the estimated net 356 students with a favorable OTI outcome, the estimated lifetime SROI is between 11,423% and 36,728%, with a total lifetime social benefit estimated between $111 million and $356 million.
- For North Carolina we do not report SROI, as none of the OTI impacts were statistically significant. The SROI values would likely represent statistical noise.

Summary and Discussion

The AT&T Aspire award to CSS represented a unique opportunity to evaluate their PGC-HS, cross-age peer mentoring and transition program. The findings reported above demonstrate that the program is working to improve short-term annual outcomes (OTI and Attendance) that are represented in both the PGC-HS and AT&T Aspire logic models. Therefore, the AT&T Aspire evaluation provides evidence that CSS has successfully implemented the program during the 2014–15 and 2015–16 academic years. This suggests that longer term outcomes (such as on-time four-year graduation and post-secondary success) might be observed as well.

Particularly impressive is the finding for both the Baltimore/NYC Project and the North Carolina Project Year 1 students see *sustained advantages for Attendance in Year 2* even when no longer participating in the PGC-HS program. This suggests the effects of the program that are gained during participation endure beyond participation in the program.

North Carolina and the Baltimore/NYC projects represent two very different community settings. The fact that CSS successfully demonstrated impacts in both regions is impressive. This finding suggests that PGC-HS may scale successfully to other contexts as well.

This evaluation found that program intensity, as measured by the number of days of student contact, made a difference in both regions. Students with more contact time had higher outcomes for OTI, GPA and Attendance in the Baltimore/NYC Project. This effect was observed for Attendance in the North Carolina Project. CSS may consider working with districts to schedule more program-student time, and possibly, design a study to identify the appropriate number of student contact hours by systematically varying program intensity across schools.

It is best if students can be compared to peers within the same district. Due to capacity and district-level challenges this was not possible for Baltimore students or students in the North Carolina Project.

This evaluation did not distinguish between mentors and mentees when examining outcomes. Also, quasi-experimental (QED) matched evaluations are not as rigorous as a randomized study can be, because QED studies can only control for observed variables. Nevertheless, this evaluation observed positive outcomes in both regions, and for both academic years. This type of replicated finding provides stronger support that PGC-HS is working as intended to positively impact participants' academic performance (OTI and GPA) and school engagement (attendance), and serves as evidence that long-term outcomes of on-time graduation and success beyond high school may be observed as well.

Appendix 1 | Sample Attrition

As the analytic process moves from the participation sample (i.e. the population of PGC-HS students), to the sample of students linked to district demographics, and finally to the impact samples, the available sample sizes become much smaller. The following figures display the relationship between the participation samples, the sample of students with district data, and the analytic sample.

The bar on the left of each figure represents the total number of students participating. The height of each subsequent bar is proportional to the participation population. Note that several impact estimates represent fewer than 12% of the original program participants. Demographics were not available for all participants however to evaluate how *representative* the impact samples are of the total participation sample. What *can* be evaluated, however, is how the closely each impact sample compares to the group of students with linked demographic data. Note that students with high mobility in and out of the district are more likely to be missing district data. Therefore students with demographic data may not be representative of all students.

The key point is that impact estimates are valid for the students included in them; but generalizing to all students in the program can be an invalid inference depending on the proportion of the original sample included, and how similar subsamples are to the population both in terms of measured and unmeasured variables.

Figure 5 | Baltimore/NYC Year 1 Sample Attrition

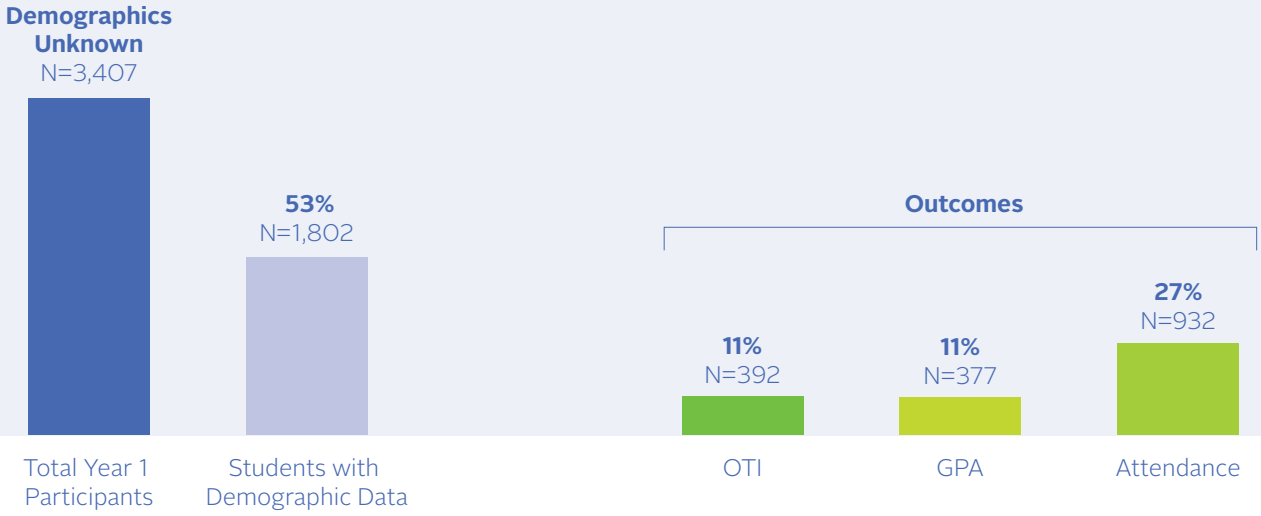


Figure 6 | Baltimore/NYC Year 2 Sample Attrition

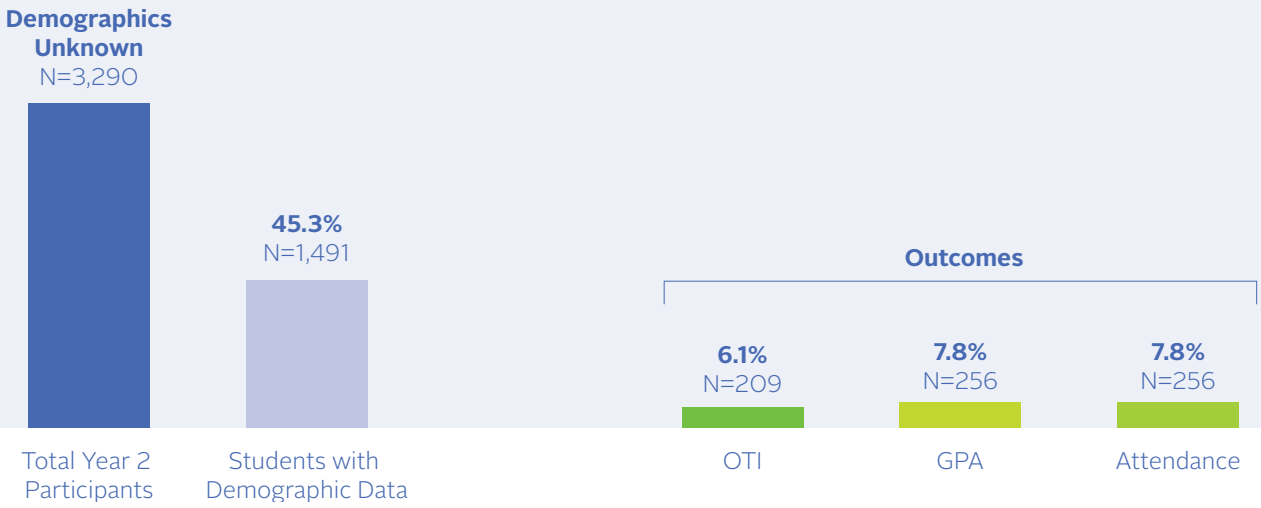


Figure 7 | Baltimore/NYC Year 1 Sustained Sample Attrition

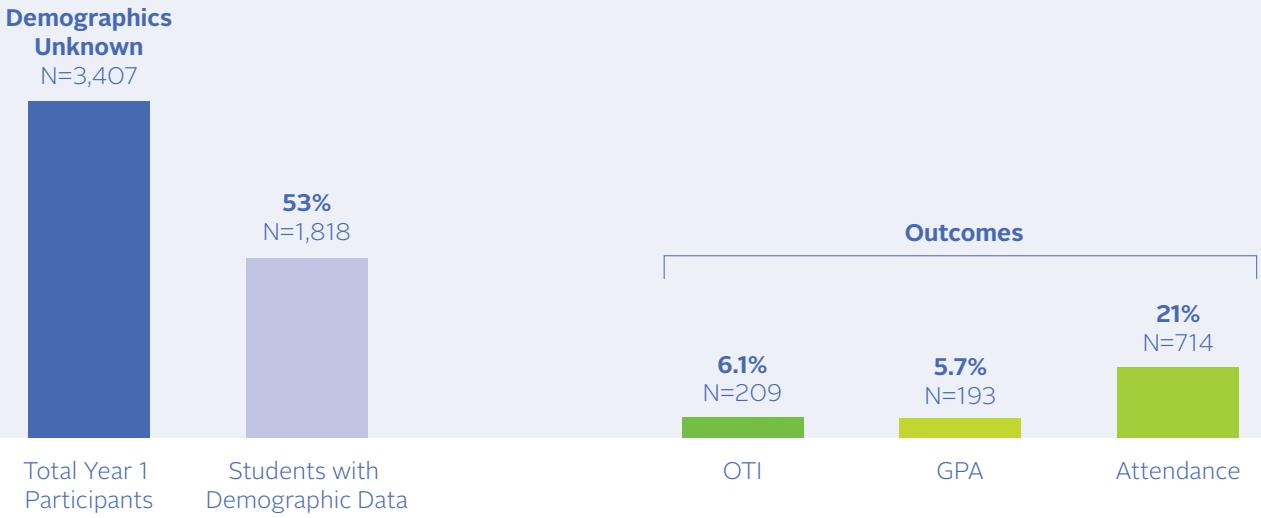


Figure 8 | North Carolina Year 1 Sample Attrition

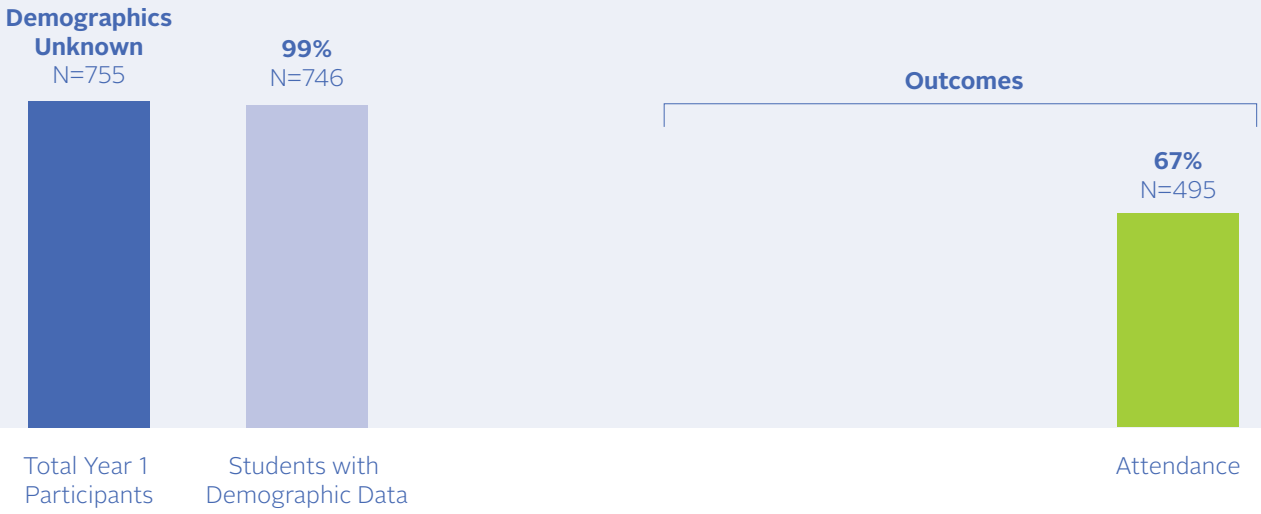


Figure 9 | North Carolina Year 2 Sample Attrition

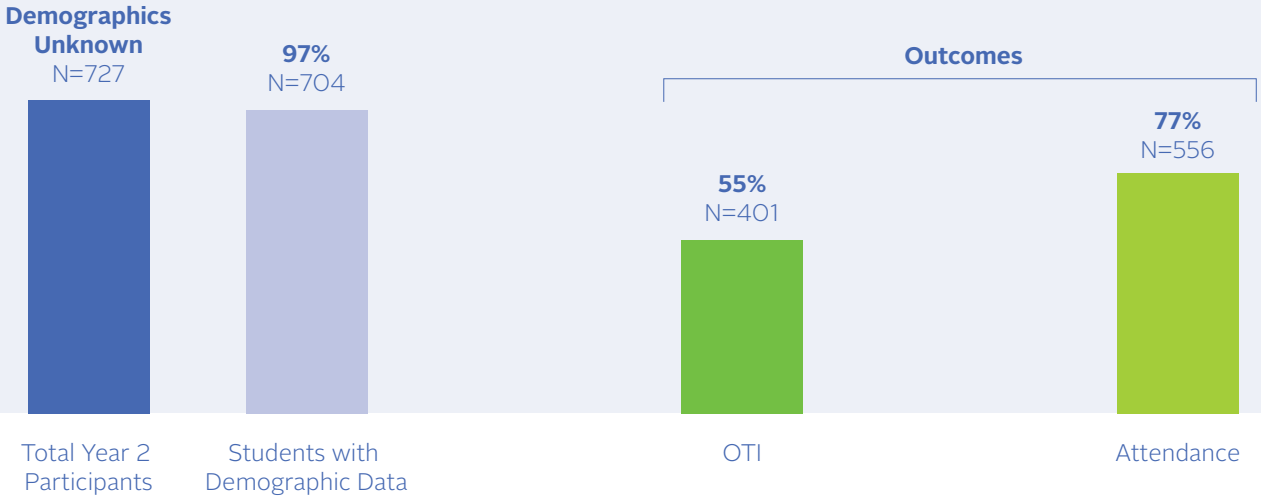
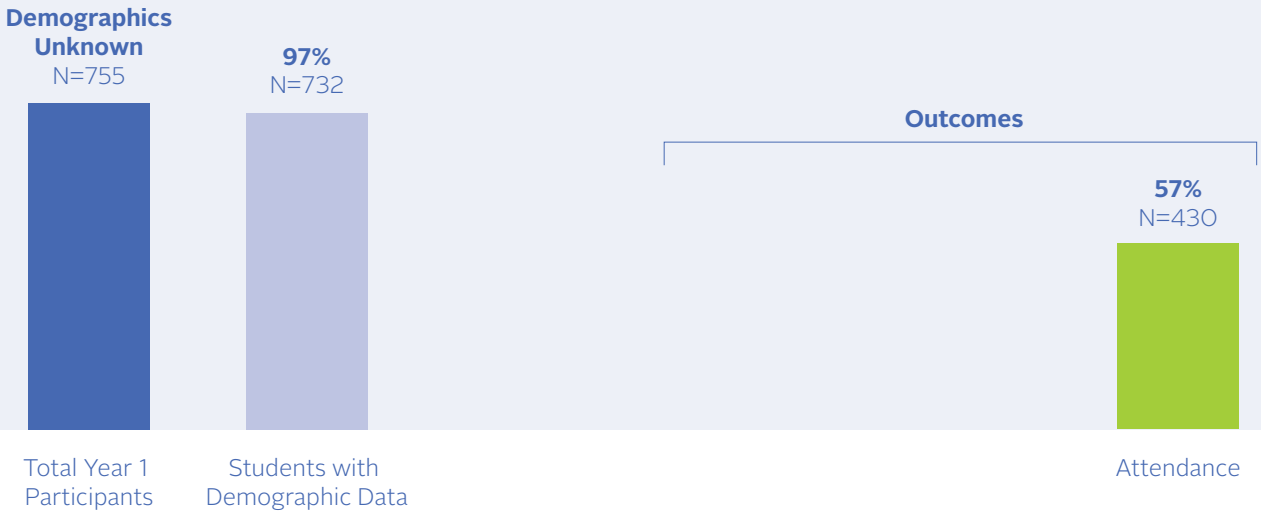


Figure 10 | North Carolina Year 1 Sustained Sample Attrition



Appendix 2 | Understanding Effect Sizes

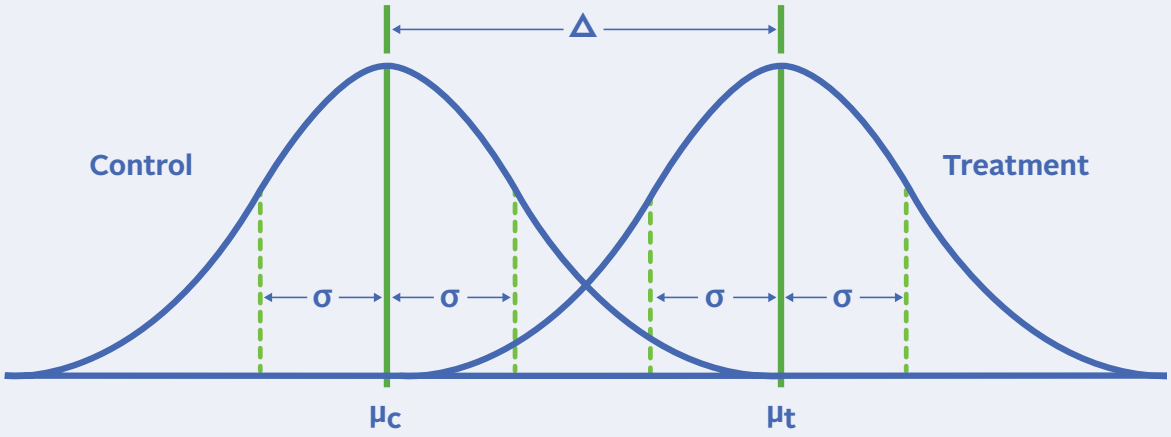
The OTI impact estimates are expressed in terms of *effect size*, or Cohen's *d*, and represent differences between two groups in standard deviation units. This allows impacts for different outcomes, or across different programs to be directly compared. An effect size of 0.50 would mean that the two groups differed by one-half of a standard deviation on the outcome.

An effect size can be understood through this simple formula:

$$ES = \frac{\bar{X}_T - \bar{X}_C}{S_p}$$

It is simply the difference in means between the program group (treatment) and the non-program group (control), divided by the pooled standard deviation of both group's data combined. This can also be visualized as two overlapping distributions of scores as in **Figure 11** below:

Figure 11 | Example of Two Distributions Revealing a Very Large Impact



Notation in this figure: μ_c and μ_t are the means of the control and treatment distribution respectively; Δ is the difference between those means, and σ is the common standard deviation of both distributions.

Note: Reproduced from Lipsey et. al. (2012).[xxiv]

What Is a Good Effect Size?

Cohen (1992)[xxv] defined effect sizes of 0.20 as *small*, 0.50 as *medium*, and 0.80 as *large*. While these standards have been widely cited, they tend to be too conservative, and don't truly reflect the range of impacts usually observed in successful educational interventions. Effect sizes of 0.10 to 0.20 are often considered educationally meaningful.

Better benchmarks than Cohen's may be offered by the typical grade-level growth on standardized assessments in high school. Bloom, Hill, Black, & Lipsey (2008) estimated that annual growth[xxvi] in grades 9-10 and 10-11 ranges from 0.14 to 0.25. The What Works Clearinghouse standards consider effects of 0.25 to be of practical significance (i.e. substantively important). Therefore, statistically significant impacts observed in this range should likely be considered good. Throughout this report we use an alpha of 0.05 with two-tailed tests to evaluate statistical significance.

Appendix 3 | AT&T Aspire Impact Estimation Procedures

Matched-Comparison Design

A matched-comparison design was used to produce an estimate of the overall effect of individual programs on participants' OTI, GPA, and school attendance. This design measures pre-existing differences between groups. However, there will always be reservations about potential bias in impact estimates due to unmeasured pre-program differences that cannot be controlled for and may be confounded (or included) in the estimate.

For the AT&T Aspire evaluation, it was necessary to categorize the level of propensity score matching (PSM) conducted. Based on the amount and quality of data provided for demographic, OTI, GPA, and school attendance, there were three levels of matching: high, medium, and low. There were two data conditions yielding high-quality matching, one data condition yielding medium-quality matching, and one data condition yielding low-quality matching. The data conditions for high-quality matching were: (1) 75 percent of program students were included in analyses, and there were data in the 2013-14 school year for at least two of the three outcomes or (2) 75 percent of program students were included in analyses, there were data in the 2013-14 school year for at least one of the three outcomes, and there were demographic data. The data conditions for medium-quality matching were: between 50 percent and 75 percent of program students were included in analysis, and there were data provided in the 2013-14 school year for at least one of the three outcomes. The data condition for low-quality matching was less than 50 percent of program students included in analyses, and there were data provided in the 2013-14 school year for at least one of the three outcomes. If there were not sufficient data to be classified into one of these three levels, students were excluded from analysis.

Steps in Propensity Matching and Quality of Matches

1. Run propensity score matching (PSM) seeking up to three matches for each program student depending on the sample size of the comparison group. Match on the natural pretest[5] of **all outcomes** and **all demographic variables** (run matching for each grade separately).

2. Run PSM seeking up to three matches for each program student depending on the sample size of the comparison group. Match on natural pretest of **all outcomes** and **a subset of demographic variables** (run matching for each grade separately).

3. Run PSM seeking up to three matches for each program student depending on the sample size of the comparison group. Match on natural pretest of all outcomes, ignoring demographics because too many are missing (run matching for each grade separately).

4. Run PSM seeking up to three matches for each program student depending on the sample size of the comparison group. Match on the natural pretest and run matching separately for each outcome, because not all the natural pretests are available (run matching for each grade separately).

It is important to note that PSM is an iterative process. If matching is not successful on the first attempt, then additional attempts at PSM are conducted. Some modifications to matching include removing covariates from the matching, adjusting the ratio of program: comparison students, or matching for each outcome separately instead of simultaneously. Concerning match quality, when the standardized difference was less than 0.25 standard deviations for the natural pretest included in PSM, the match was considered to be of high quality. When the standardized difference was greater than 0.25 standard deviations for the covariates included in the model, the match was considered unacceptable.

[5] The natural pretest for any outcome, is simply that outcome from the prior year (or before the start of the intervention). Therefore, the natural pretest for the GPA outcome in 2014-15, is simply GPA from 2013-14.

The use of PSM with pre-program covariates, a subset of covariates, and the criterion for evaluating the quality of the match reflect a rigorous methodological approach to ensure bias is not introduced by running PSM when its assumptions are violated, and when the data were not adequate for PSM.

When reading the results, the sample sizes for estimates of individual recipients' effects, with reservations, will vary by student grade level (9th or 10th, 11th, and 12th grade) and by outcome (OTI, GPA, and school attendance). The following describe the specific methods and models used to estimate overall effects of an AT&T Aspire program. The description is organized by design and outcome, starting with the primary outcome of OTI and followed by school attendance and GPA.

AT&T Aspire Recipient Effect and Estimation Procedures

Effect of Individual AT&T Aspire Programs on OTI

A single-level regression model was used to estimate an individual recipient's effect on participants' outcomes. The estimate is defined as the difference between the program group and non-program group on OTI with covariate adjustments for students' pre-program characteristics. The purpose of the covariates was to statistically control for pre-existing program differences between the two groups. The design relies on matching to create equivalent groups, and any residual differences are statistically controlled for by the covariates.

Single-level regression models were used to estimate the effect of the individual recipients on participants' OTI as the primary, dichotomous outcome and school attendance and GPA as the secondary outcomes. To estimate the individual recipient's effect, the linking function for the general linear model is defined as follows:

Let Y=1=on track to graduate in 2015-16 (success);

Let Y=0=off track to graduate in 2015-16 (failure).

Let $\pi(Y=1 \mid X_1, X_2, \ldots X_q) = \pi(x)$ represents the probability of success or the probability of being on track to graduate on time for a given set of *q* explanatory variables.

The logistic model for estimating the log-odds of being on track to graduate on time is as follows:

$$\text{Ln}(Y')=\text{logit}[\pi(x)=\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = B_0 + B_1\,(\text{Recipient}_{i1}) + B2(\text{PreCov}_{i2}) + \ldots + B_k(\text{PreCov}_{ik})$$

where,

$Recipient_{i1}$ is the ith student in the sample with a value of 1=enrolled in an individual recipient's Aspire program or 0=not enrolled in an individual recipient's AT&T Aspire program.

$PreCov_{i1}$ represents the *i*th student in the sample with a value on a pre-program covariate entered into the model as a statistical control. For example, if the outcome is OTI16 then the value PreCov_{i2} is 1=on track to graduate or 0=off track to graduate prior to individual recipient's enrollment if *i*th student in the sample has a value of 1 on $Recipient_{i1}$, or prior to individual recipient's delivery if *i*th student in the sample has a value of 0 on $Recipient_{i1}$.

B_1 is the log-odds of the *i*th student enrolled in the individual recipient's AT&T Aspire program being on track to graduate relative to the *i*th matched-comparison student not enrolled in the individual recipient's AT&T Aspire program. The standard error and *p* value can be obtained from the regression output. The log-odds is typically interpreted in terms of the odds ratio that can be obtained by taking the exponent of B_1 in the form of $\exp^{(B1)}$.

$PreCov_{ik}$ represents the *i*th student in the sample with a value on the *k*th pre-program covariate entered into the model as a statistical control.

B_2 is the log-odds of the ith student enrolled in the individual recipient's AT&T Aspire program being on track to graduate relative to the *i*th matched-comparison student not enrolled in the individual recipient's AT&T Aspire program. The standard error and *p* value can be obtained from the regression output. The log-odds is typically interpreted in terms of the odds ratio that can be obtained by taking the exponent of B_k in the form of $\exp^{(Bk)}$.

Effect of Individual AT&T Aspire Programs on Attendance and GPA

The school attendance outcome (Attendance16), which is conceptualized in this evaluation as a measure of school engagement, and GPA16 were both measured on a continuous scale. An ordinary least squares regression model was used that mirrored the regression model used to estimate the individual recipient's effect on OTI, except the logit link was replaced by the continuous outcome measured in percentage of school days attended or GPA.

$$Y_i = B_0 + B_1 (Recipient_{i1}) + B_2(PreCov_{i2}) + \ldots + B_k(PreCov_{ik}) + e_i$$

where,

Y_i is the 2014-15 school attendance or 2014-15 academic GPA for the ith student in the analysis sample,

$Recipient_{i1}$ is the ith student in the sample with a value of 1=enrolled in individual recipient's AT&T Aspire program or 0=not enrolled in individual recipient's AT&T Aspire program.

$PreCov_{i1}$ represents the *i*th student in the sample with a value on a pre-program covariate entered into the model as a statistical control. For example, if the outcome is OTI, then the value $PreCov_{i2}$ is 1=on track to graduate or 0=off track to graduate prior to individual recipient's program enrollment if *i*th student in the sample has a value of 1 on $Recipient_{i1}$, or prior to individual recipient's program delivery if *i*th student in the sample has a value of 0 on $Recipient_{i1}$.

B_1 is the average difference between the program and non-program group in days of absences (or GPA) adjusted for covariates in the model.

$PreCov_{ik}$ represents the *i*th student in the sample with a value on the *k*th pre-program covariate entered into the model as a statistical control.

B2 is the one-unit change in the outcome for a one-unit change on the kth covariate for the *i*th student in the sample.

Matching Attempt Logic

First matching attempt

1. Matching variable
   - **a.** All natural baselines (atr_1415, oti_1415 (or cred_1415 & yfail_1415), gpa_1415
     - *i.* If all natural baseline variables and outcome variables have a similar missing rate, this option makes sense. If there is a substantial missing rate difference, we may need to match one outcome and its pretest at a time.
   - **b.** All demographic variables
2. Matching ratio
   - **a.** 1:3
     - *i.* If comparison pool:program ratio is small and the sample size is sufficiently large to exceed 50, the ratio can be reduced.
     - *ii.* If comparison pool:program ratio is large and program sample size is small, the ratio can be increased.
3. Success criteria
   - **a.** SMDs for all matched variables are below 0.25.
     - *i.* This criteria can be relaxed (*e.g., up to three demographic variables exceeding 0.25 is OK, while all pretests must be below 0.25.*)
   - **b.** Data reduction rate due to matching is below 25 percent.
     - *i.* Data reduction rate due to include three natural pretests also needs to be considered.
   - **c.** Total sample size is 50 or greater.

Second matching attempt

1. Matching variable
   - **a.** All natural baselines.
   - **b.** Drop some demographic variables (how many demographic variables need to be kept?).
2. Matching ratio
   - **a.** Reduce if comparison pool:program ratio is small.
   - **b.** Increase if total matched sample size was below 50.

3 Success criteria

- **a.** SMDs for all natural pretests and at least one demographic variable are below 0.25.
- **b.** Data reduction rate due to matching is below 25 percent.
- **c.** Total sample size is 50 or greater.

Third matching attempt

1 Matching variable

- **a.** All natural pretest.

2 Matching ratio

- **a.** Reduce if comparison pool:program ratio is small.
- **b.** Increase if total matched sample size was below 50.

3 Success criteria

- **a.** SMDs for all natural pretests are below 0.25.
- **b.** Data reduction rate due to matching is below 25 percent.

Total sample size is 50 or greater.

Fourth and more attempts

1 Match one outcome at a time.

2 Replicate first, second, and third attempts with single outcome and natural baseline.

Appendix 4 | Social Return on Investment

The first step in social return on investment (SROI) is estimating a program's impact on program participants' outcome (OTI). The results from this evaluation were the basis for all ROI calculations.

A favorable OTI outcome for any AT&T Aspire program can be achieved when a student has moved from off-track to on-track status or has been prevented from going off track during the academic year.

$$O_i = [(OTI_T^+ - OTI_C^+) - (OTI_T^- - OTI_C^-)]$$

where O_i is the net number of students with favorable outcomes, which is the number of students with favorable outcomes minus the number of students with unfavorable outcomes; OTI_T^+ is the number of students in the program moving from off track to on track; OTI_C^+ is the number of students in the control group moving from off track to on track; OTI_T^- is the number of students in the program moving from on track to off track; and OTI_C^- is number of students in the control group moving from on track to off track.

In the specific example below, the net number of students with favorable outcomes for the program equals three (O_1=3), which is the number of students with favorable outcomes minus the number of students with unfavorable outcomes, where the number of students in the program group moving from off track to on track is ten ($OTI_T^+ = 10$); the number of students in the control group moving from off track to on track is five ($OTI_C^+ = 5$); the number of students in the program moving from on track to off track is eight ($OTI_T^- = 8$); and the number of students in the control group moving from on track to off track is six ($OTI_C^- = 6$).

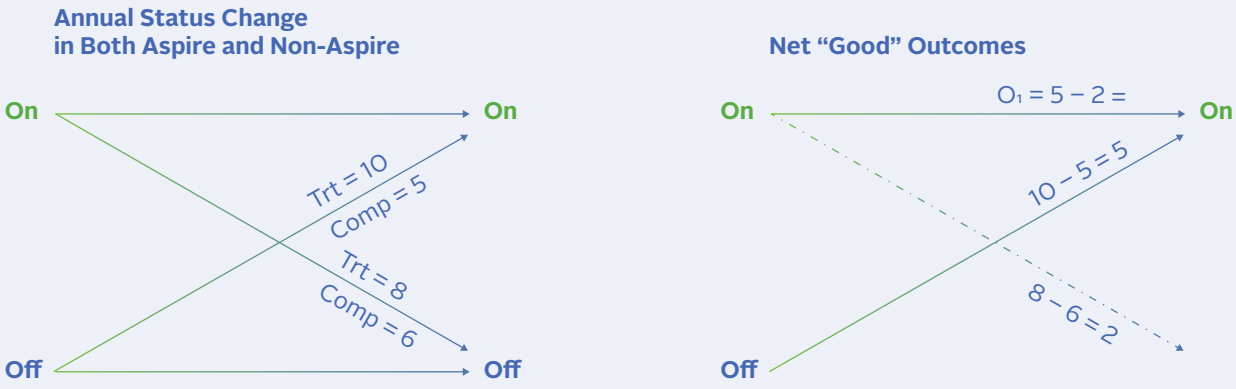
$$O_1 = [(10 - 5) - (8 - 6)] = 3$$

ROI Analysis

The SROI has been calculated using the classic ROI formula (New Education Foundation, 2004):

$$SROI = \frac{Benefit\ from\ investment - Cost\ of\ investment}{Cost\ of\ investment} \times 100\%$$

Figure 12 | OTI Net Mover Analysis Example



The benefit from investment was estimated based on the forecasted lifetime income of each individual (Table 10). The data for this estimation were retrieved from Education and Synthetic Work-Life Earnings Estimates, a study conducted by the U.S. Census Bureau (Julian & Kominski, 2011).

Table 10 | Social Benefit Estimation for Each Student

	Optimistic	Conservative
Work-life earning before high school graduation	$1,099,000	$1,099,000
Work-life earning after high school	$1,968,594	$1,371,000
Increase in work-life earning after high school	$869,594	$272,000
Tax rate (fixed)	15%	15%
Increase in tax payment	$130,439	$40,800
Total social benefit	**$1,000,033**	**$312,800**

The benefit of an AT&T Aspire program is calculated by multiplying the number of students with favorable outcome results and the estimated social benefit for each AT&T Aspire student. The following formula shows the calculation:

$$O_i = [(OTI_T^+ - OTI_C^+) - (OTI_T^- - OTI_C^-)]$$

$$B_P = O_i \times B$$

B_P is the social benefit of each AT&T Aspire program; and B is the social benefit for each AT&T Aspire student.

References

[i] http://gradnation.americaspromise.org/

[ii] https://consortium.uchicago.edu/publications/track-indicator-predictor-high-school-graduation

[iii] Blum, R. (2005). *School Connectedness: Improving the Lives of Students*. Baltimore, MD: Johns Hopkins Bloomberg School of Public Health.

[iv] Fredricks, J. Blumenfeld, P. & Paris, A. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*, 59-109.

[v] Bridgeland, J.M., DiIulio, J.J., Morison, K.B. (2006). *The Silent Epidemic: Perspectives of High School Dropouts*. Washington, D.C.: Civic Enterprises and Peter D. Hart Research Associates.

[vi] Cohen, J., & Smerdon, B. (2009). Tightening the dropout tourniquet: *Easing the transition from middle to high school. Preventing School Failure, 53*, 177-183.

[vii] Kennelly, L., & Monrad, M. (2007). *Easing the Transition to High School: Research and Best Practices Designed to Support High School Learning*. Retrieved from http://www.betterhighschools.org/docs/NHSC_TransitionsReport.pdf

[viii] Wheelock, A., & Miao, J. (2005). The ninth-grade bottleneck: An enrollment bulge in a transition year that demands careful attention and action. *The School Administrator, 62*(3), 36.

[ix] Yeager, D.S. (2017). Social and emotional learning programs for adolescents. *The Future of Children, 27* (1), 73-94.

[x] Yeager, D.S. (2017). Social and emotional learning programs for adolescents. *The Future of Children, 27* (1), 73-94.

[xi] Yeager, D.S. (2017). Social and emotional learning programs for adolescents. *The Future of Children, 27* (1), 73-94.

[xii] https://greatergood.berkeley.edu/article/item/why_dont_students_take_social_emotional_learning_home

[xiii] https://greatergood.berkeley.edu/article/item/why_dont_students_take_social_emotional_learning_home

[xiv] Tough, P. (2012). *How Children succeed. Grit, curiosity, and the hidden power of character*. New York, New York: Houghton Mifflin Harcourt Publishing Company.

[xv] Collaborative for Academic, Social, and Emotional Learning (CASEL). (2007). *What is SEL? Skills and competencies*. Chicago, IL: Author.

[xvi] Wellman, B. (1988). Structural analysis: From method and metaphor to theory and substance. In B. Wellman B & S.D. Berkowitz (Eds.), *Social structures: A network approach* (pp.19-61). Cambridge, MA: Cambridge University Press.

[xvii] Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

[xviii] McCormick, M. P. & Cappella, E. (2014). Conceptualizing academic norms in middle school: A social network perspective. *Journal of Early Adolescence*, advanced online publication. doi: 10.1177/0272431614535093

[xix] Wellman, B. (1988). Structural analysis: From method and metaphor to theory and substance. In B. Wellman B & S.D. Berkowitz (Eds.), *Social structures: A network approach* (pp.19-61). Cambridge, MA: Cambridge University Press.

[xx] Cappella, E., & Hwang, S. (in press). Peer contexts in schools: Avenues toward behavioral health in early adolescence. *Behavioral Medicine*.

[xxi] Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

[xxii] Isakson, K., & Jarvis, P. (1999). The adjustment of adolescents during the transition to high school: A short-term longitudinal study. *Journal of Youth and Adolescence, 28*(1), 1-26.

[xxiii] Julian & Kominski, (2011). The data for this estimation were retrieved from Education and Synthetic Work-Life Earnings Estimates, a study conducted by the U.S. Census Bureau.

[xxiv] Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education. This report is available on the IES website at http://ies.ed.gov/ncser/.

[xxv] Cohen, J (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.

[xxvi] Bloom, H. S., Hill, C. J., Black, A. B., and Lipsey, M. W. (2008). Performance Trajectories and Performance Gaps as Achievement Effect-Size Benchmarks for Educational Interventions. *Journal of Research on Educational Effectiveness*, 1(4): 289-328.