Mathematica®
Progress Together

# Air Tutors' Online Tutoring: Math Knowledge Impacts and Participant Math Perceptions

**Lily Fesler, Anna Gu, and Greg Chojnacki**

**March 2023**

## Contents

Executive Summary ..... iv

Introduction ..... 1

Key Findings ..... 1

Air Tutors students achieved higher scores on the MAP assessment than control group students. ..... 1

Students who participated in Air Tutors reported positive relationships with their tutors, a high sense of belonging in their tutoring sessions, and increased math confidence over the course of the program ..... 2

Student attendance was higher at sites where tutoring took place during the school day and lower at sites where tutoring occurred after school ..... 3

Overarching Conclusion and Next Steps ..... 4

References ..... 8

## Tables

Table 1. Estimated impacts of Air Tutors on spring MAP assessment scores ..... 2

Table B.1. Number of treatment and control group students who were randomly assigned and who have outcome data, by site ..... 10

Table B.2. Characteristics of treatment and control group students at baseline ..... 11

## Figure

1. Percentage of student survey responses that were at least 4 (out of 5) ..... 3

# Middle Years Math Grantee Report Series

This report is one in a series of six reports on math tutoring programs. Over the 2020–2021 and 2021–2022 school years, the Bill & Melinda Gates Foundation invested in rapid-cycle evaluations of a cohort of 10 tutoring providers to learn about their innovative approaches to tutoring as part of its Middle Years Math body of work.[1] The goal of these investments was to understand how different tutoring models might create positive student experiences and lead to improved academic outcomes for students in the foundation's priority communities—those who are Black, Latino, and/or experiencing poverty. These investments were grounded in the substantial body of evidence supporting the effectiveness of tutoring in improving student math knowledge (Nickow et al., 2020).

To build on this existing evidence of effectiveness, the Gates Foundation sought to develop new early evidence about the success of a range of tutoring approaches. Specifically, these investments targeted two key learning priorities. First, the foundation sought to learn how innovative technologies and tutoring program design features might simultaneously improve the quality and lower the cost of tutoring, making high-quality tutoring available to a large number of students in priority communities. The second priority was to learn the extent to which tutoring programs resulted in positive experiences for participating students. To learn about tutoring design features, the foundation invested in tutoring programs with a wide range of approaches, including group and one-on-one tutoring, virtual and in-person models, professional teachers as tutors, or volunteer tutors who shared aspects of identity with tutored students. Tutoring programs also used different approaches to tutoring curriculum and pedagogy. The goal of this report series is to inform the tutoring field more broadly and support the provision of high-quality tutoring to as many students in the priority communities as possible.

To learn rapidly about tutoring providers' innovative approaches, Mathematica worked with each one to identify the most rigorous study design that would be feasible for district partners within a one-to-three-month planning period. Some providers were able to design and implement randomized controlled trials; others used quasi-experimental designs such as matched comparison approaches. One study compared growth in math knowledge among participants to the growth observed in national samples because it was not possible to obtain student-level data for comparison students who did not receive tutoring. These relatively small studies were right-sized to the development stage of the tutoring program and sought to demonstrate early evidence of success before moving on to larger-scale effectiveness studies. To help synthesize findings about student experiences from multiple providers, studies used the same student survey measures of tutor relationship, math confidence, and sense of belonging in tutoring sessions. Most of the studies used standardized math knowledge assessments aligned with Common Core State Standards.

Each study also aimed to inform providers' efforts to refine their programs and support successful implementation. These studies measured the amount of tutoring offered, attendance, and staff impressions about implementation challenges while also gathering qualitative data on students' experiences. Findings from these studies have helped to direct tutoring providers' next steps in refining and scaling their tutoring programs.

[1] This publication is based on research funded by the Bill & Melinda Gates Foundation. The findings and conclusions contained within are those of the authors and do not necessarily reflect positions or policies of the Bill & Melinda Gates Foundation.

# Executive Summary

## What is the tutoring program we studied?

Air Tutors provides live online tutoring on math, English, and reading to K–12 students, both one-on-one and in small groups, using an online platform that incorporates video conferencing and interactive whiteboards. Its program focuses on providing high-quality paid tutors with prior tutoring experience and engaging online personalities, as well as on engaging students and their families via text messages, phone calls, and emails. The program recruits tutors who hold at least a bachelor's degree and three years of tutoring or teaching experience, with a preference for tutors who hold at least a master's or Ph.D. with five or more years of tutoring or teaching experience. Additional details about Air Tutors are available at https://airtutors.org and in Appendix A. This study focused on Air Tutors' math tutoring program for students in grades 4 through 6 in small groups of three and four students. Air Tutors' math tutoring program provides instruction on foundational math skills, including addition, subtraction, decimals, fractions, proportions, and basic geometry.

## What questions does this study answer?

1. What effects does Air Tutors have on participating students' math knowledge, as measured by performance on standardized math tests?
2. Do Air Tutors students report higher levels of math confidence after receiving tutoring than before?
3. Do Air Tutors students report a high sense of belonging in their tutoring sessions?
4. Do Air Tutors students report positive student–tutor relationships?
5. What is the average attendance rate for participating Air Tutors students?

## How was the study conducted?

**Study design.** The 2021–2022 program took place across three schools, including a K-12 online school (District 1) and an elementary school and middle school in a school district in Texas (District 2). Forty-seven 4th-, 5th-, and 6th-grade students participated in the program in District 1 out of 64 who were invited, and 106 4th-, 5th-, and 6th-grade students participated in the program in District 2 out of 144 who were invited. Air Tutors students were randomly selected from eligible students at each school. Eligible students included those who were available during tutoring times (for both districts), were not receiving other additional services (District 1), did not have Individualized Education Programs (District 1 and the middle school in District 2), and were up to three grade levels below average on the Measures of Academic Progress (MAP) assessment (District 2). Students in the control group were those who were eligible for tutoring but were not selected to participate in tutoring. The program offered students a total of three hours of tutoring over three or four sessions each week, for a total of 70 hours of tutoring in District 1 and 42 hours of tutoring in District 2. For 6th graders in District 2 and all students in District 1, sessions took place during the school day. For 4th- and 5th-grade students in District 2, sessions occurred after school. Sessions for students in District 1

spanned both semesters, whereas sessions for District 2 students took place only in spring 2022.

**Methods.** This study used a stratified randomized controlled trial to compare the math knowledge outcomes (measured by the MAP assessment) of students participating in Air Tutors versus students in the control group. We stratified the sample based on students' school, grade, and the class period they had available for tutoring to ensure similar numbers of students from each of these groups would be served. We examined the impact of Air Tutors using a linear regression model, which controlled for the strata (based on the three aforementioned characteristics), students' baseline math scores, and student demographic characteristics. The study also collected survey data from Air Tutors students before, during, and after the study period: the Patterns of Adaptive Learning Scale (PALS) and the Math and Me survey to measure confidence in math, and the Copilot-Elevate survey to measure the student–tutor relationship and sense of belonging in tutoring sessions and in their math class. We used descriptive quantitative analysis of student survey data, as well as attendance data of tutoring sessions. The sample included 276 students with complete math achievement data (including treatment and control group students) and 105 participating treatment students with survey data at both baseline and at the end of the year. (The survey was not administered to control group students.)

**Limitations.** The main limitations of this study relate to the representativeness of the sample of students. First, the findings only reflect students who have outcome and survey data. Twelve percent of randomly assigned students did not have outcome data on the MAP assessment, and 31 percent of Air Tutors participants did not take either the baseline or end-of-year survey. Second, the main impacts presented in this report only apply to the 74 percent of students who chose to participate in the program, and impacts may differ for students who would not choose to voluntarily participate in the program. Lastly, students were not randomly assigned to receive tutoring at different times or in different school settings, so any differences in impacts associated with these contextual factors could be due to other factors as well.

## What did the study find?

**Math knowledge.** MAP scores among students participating in Air Tutors were 0.13 standard deviations higher than those among control group students, after accounting for baseline differences. This is equal to a 4.0 percentile point increase from students who did not participate in Air Tutors. A Bayesian analysis, which provides a probability that the intervention had an effect given the impact estimate and prior evidence on similar interventions, indicated that Air Tutors had a 93 percent chance of boosting math knowledge. Air Tutors increased participating students' MAP scores the most at a program held during the school day at an in-person middle school (0.26 standard deviations). The analysis also showed that Air Tutors had a 66 percent chance of increasing students' MAP scores at an online school where tutoring was held during the school day (with an estimate of 0.10 standard deviations) and a 75 percent chance of increasing students' MAP scores at an elementary school where tutoring was held as an after-school program (with an estimate of 0.08 standard deviations).

**Student survey.** Students participating in Air Tutors reported highly positive relationships with their tutors, with 89 percent of District 1 students and 87 percent of District 2 students rating an average of at least 4 on a 5-point scale at the end of the year, representing agreement or strong agreement with statements about positive relationships. They also reported a high sense of belonging in their tutoring sessions, with 80 percent of students in District 1 and 69 percent of students in District 2 rating an average of at least 4 on a 5-point scale at end-of-year, representing agreement or strong agreement with statements about sense of belonging. Finally, student math confidence increased over the course of the program, with students from both districts reporting an average increase of 0.18 points on a 5-point scale from baseline to the end of the year.

**Tutoring session attendance.** On average, students attended 72 percent of sessions. The attendance rate was 63 percent at the in-person elementary school, where the virtual tutoring was offered after school at the school's campus. The attendance rates were higher at the in-person middle school (78 percent) and the online school (79 percent), both of which offered virtual tutoring during the school day.

## Introduction

Air Tutors is an online tutoring organization that provides one-on-one and small group tutoring for K–12 students. Air Tutors uses live, online tutoring and an online platform that incorporates video conferencing and interactive whiteboards; recruits high-quality paid tutors with tutoring experience and engaging online personalities; and engages students and their families via text messages, phone calls, and emails.

This study aims to provide evidence of the impact of tutoring on student math achievement in a randomized controlled trial, as well as on students' self-reported math confidence, sense of belonging, and relationships with their tutors via student surveys. Specifically, we examined the impact of participating in Air Tutors math tutoring sessions on 4th-, 5th-, and 6th-grade students' math achievement, as well as changes in student math confidence. In this evaluation, we sought to answer the following research questions:

1. What effects does Air Tutors have on have on participating students' math knowledge, as measured by performance on standardized math tests?
2. Do Air Tutors students report higher levels of math confidence after receiving tutoring than before?
3. Do Air Tutors students report a high sense of belonging in their tutoring sessions?
4. Do Air Tutors students report positive student–tutor relationships?
5. What is the average attendance rate for participating Air Tutors students?

## Key Findings

### Air Tutors students achieved higher scores on the MAP assessment than control group students.

Spring MAP assessment scores among Air Tutors students from both districts who participated in at least one Air Tutors session ("participating students") were 0.13 standard deviations higher than those of control group students, although this difference was not statistically significant ($p = 0.12$, Table 1). A Bayesian analysis indicated that Air Tutors had a 93 percent chance of boosting students' math knowledge. (See Appendix B for details on analysis methods.)

Seventy-four percent of students who were invited to participate in Air Tutors participated in the program (defined as attending at least one Air Tutors session). The impact of Air Tutors on the full sample of invited students, regardless of whether they participated, was 0.10 standard deviations. Both impact analyses accounted for students' baseline characteristics, although there were no significant differences between treatment and control group students at baseline (see Table B.2 in Appendix B).

Air Tutors increased participating students' MAP scores in the District 2 middle school by 0.26 standard deviations, relative to control group students, and this difference was statistically significant at the 0.05 level. This was a somewhat larger impact than at the other two schools, where participating Air Tutors students' MAP scores were higher than control group students' MAP scores (by 0.10 standard deviations at the District 1 online school and by 0.08 standard deviations at the District 2 elementary school). A Bayesian analysis showed that Air Tutors had a 66 percent chance

of increasing students' MAP scores at the District 1 online school and a 75 percent chance of increasing students' MAP scores at the District 2 elementary school.

**Table 1. Estimated impacts of Air Tutors on spring MAP assessment scores**

| | Participating students | Invited students |
|---|---|---|
| All sites | 0.13<br>(0.09) | 0.10<br>(0.06) |
| District 1 online school | 0.10<br>(0.19) | 0.07<br>(0.14) |
| District 2 elementary school | 0.08<br>(0.14) | 0.05<br>(0.09) |
| District 2 middle school | 0.26**<br>(0.12) | 0.26<br>(0.15) |

Source: District data.

Note: The values in parentheses below each impact estimate are standard errors. Both sets of impact estimates (impacts among participating and invited students) use information from all 276 evaluation sample members, but the estimated impact among participating students applies to the 143 students who participated in at least one tutoring session, whereas the impact among invited students applies to all 184 students who were invited. See Table B.1 in Appendix B for sample sizes in the MAP analyses by site.

* Significantly different from zero at the .10 level, two-tailed test.

** Significantly different from zero at the .05 level, two-tailed test.

*** Significantly different from zero at the .01 level, two-tailed test.

### Students who participated in Air Tutors reported positive relationships with their tutors, a high sense of belonging in their tutoring sessions, and increased math confidence over the course of the program.

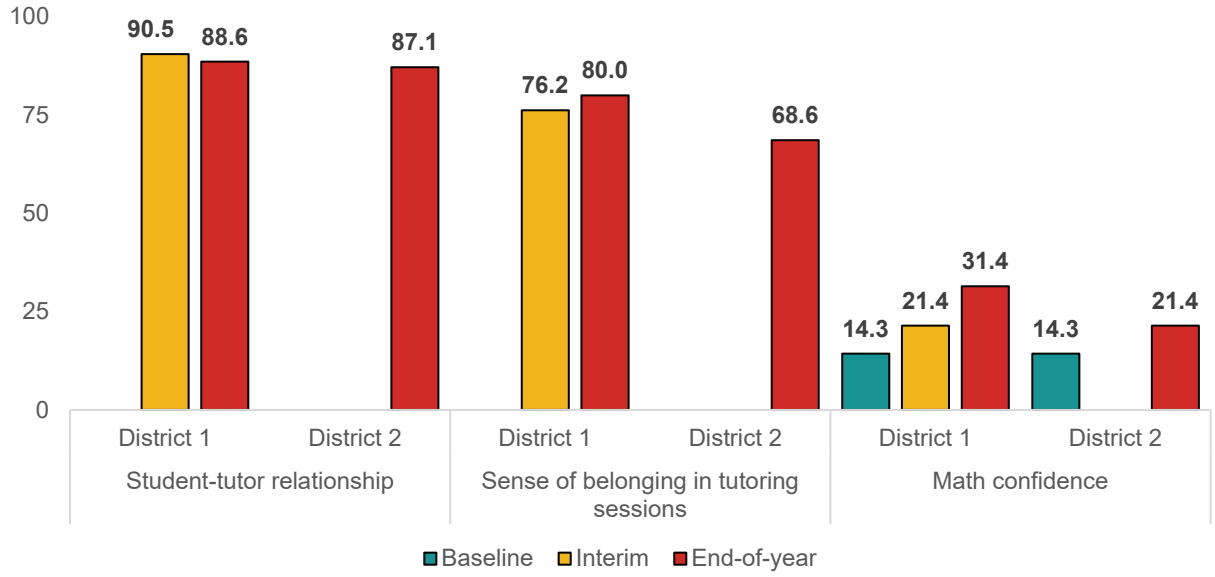
A majority of students who participated in Air Tutors reported highly positive relationships with their tutors. At the end of the year, 88 percent of students agreed or strongly agreed with three statements about whether their tutor treats them with respect, is glad that they are in the tutor's class, and cares about what they think (Figure 1). Students reported similarly positive relationships with their tutors across the District 1 online school interim survey and the District 1 and District 2 end-of-year surveys.

Participating students reported a high sense of belonging in their tutoring sessions, with 80 percent of District 1 students and 69 percent of District 2 students agreeing or strongly agreeing with statements about this outcome on the end-of-year survey. Sense of belonging survey items included statements about whether students had the opportunity to get to know their classmates better, whether students feel like their tutor accepts them for who they are as a person, and whether they feel comfortable sharing their thoughts and opinions during tutoring sessions.

Participating students' self-reported math confidence increased over the course of the program. Fourteen percent of participating students reported high levels of math confidence at the beginning of the program (defined as an average response of 4 or greater on a 5-point scale across 12 related survey items), and 25 percent of participating students reported high levels of math confidence at the end of the program. This increase was statistically significant at the 0.01 level. Items on the math confidence survey included statements about whether students are certain that they can learn everything taught in math and whether students feel that they are really good at math.

At the end of the program, one-half to two-thirds of participating students also reported that they felt a high sense of belonging in math class (65 percent of students), had received feedback for growth in tutoring sessions (63 percent of students), and had a voice in tutoring sessions (55 percent of students). The topic of feedback for growth included statements such as whether students received specific suggestions on how to improve their skills; sense of belonging included statements such as whether students felt comfortable sharing their thoughts and opinions in Air Tutors sessions and in their math class; and student voice included statements such as whether students had the opportunity to make a choice about their work.

**Figure 1. Percentage of student survey responses that were at least 4 (out of 5)**




Source: Online survey data collected by Air Tutors.

Note: The baseline and end-of-year results include 35 District 1 students and 70 District 2 students who responded to both surveys. The interim results include 42 District 1 students who responded to the baseline and interim surveys. The results are similar (within 2 percentage points) for students in District 1 who responded to all three surveys. See Appendix B for more information on the survey. Questions about student–tutor relationship and sense of belonging in tutoring sessions were not included on the baseline survey. District 2 did not administer interim surveys because its tutoring program was only one semester.

### Student attendance was higher at sites where tutoring took place during the school day and lower at sites where tutoring occurred after school.

Student attendance in tutoring sessions overall was 72 percent. Attendance was highest at the District 1 online school (79 percent) and the District 2 middle school (78 percent), where students received tutoring during the school day, and was lowest at the District 2 elementary school (63 percent), where students received tutoring after school. The offer of tutoring during the school day for 6th graders in District 2 likely explains their higher attendance, compared to 4th and 5th graders in District 2 who were offered tutoring after school. The higher attendance rates in the middle school relative to the elementary school may partially explain the higher estimated impact in the middle school compared to the elementary school in District 2.

The director of after-school programs at the District 2 elementary school reported that students would sometimes miss Air Tutors' after-school program if they lacked after-school transportation,

during inclement weather, or due to being sick. Students in the District 2 middle school also reported sometimes missing Air Tutors due to being unable to get to school on time, being pulled out of class (for example, for counselor meetings or doctor appointments), or being sick, but students in the middle school did not miss sessions as frequently as students in the elementary school.

## Overarching Conclusion and Next Steps

Study findings from District 2, where students attended school in person rather than virtually, suggest that students attended more often and may have increased their math knowledge more in the Air Tutors program that took place during the school day (the District 2 middle school) compared to the program that took place after school (the District 2 elementary school.) This is likely due, at least partially, to increased implementation challenges in the after-school setting. The middle school in District 2, in addition to Air Tutors staff, typically took attendance during tutoring sessions, which may have also provided a sense of increased accountability. Air Tutors will try to hold future programs during the school day to avoid the after-school implementation challenges or will provide additional supports for tutoring programs if they must occur after school.

Air Tutors reported that many students in the District 1 online school had adult guardians or parents available at home to help support students in their sessions, which made tutoring run fairly smoothly. Air Tutors staff also were able to easily build rapport with students at this school because the students were accustomed to remote interactions at their online school. Future implementations might include other online schools where students have support from guardians or parents overseeing their progress.

Air Tutors will also continue to focus on both maintaining direct lines of communication with tutors and building strong partnerships with teachers in future programs. In this program, Air Tutors employed staff who were responsible for responding to tutor feedback about issues during tutoring sessions and who built a line of communication with teachers at partnering schools. This provided clarity and shared understanding on issues that arose, such as attendance and discipline. Air Tutors also found that holding site presentations to inform teachers and counselors about the program in District 2 was an effective strategy and they expect to continue to hold these types of presentations in future programs.

## References

Adelson, J. L., & McCoach, D. B. (2011). Development and psychometric properties of the Math and Me survey: Measuring third through sixth graders' attitudes toward mathematics. *Measurement and Evaluation in Counseling and Development, 44*(4), 225–247.

Deke, J., Finucane, M., & Thal, D. (2022). *The BASIE (BAyeSian Interpretation of Estimates) framework for interpreting findings from impact evaluations: A practical guide for education researchers* (NCEE 2022-005). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/pubs/2022005/

Midgley, C., Maehr, M., Hruda, L., Anderman, A., Anderman, L., Freeman, K., Gheen, M., Kaplan, A., Kumar, R., Middleton, M., Nelson, J., Roeser, R., & Urdan, T. (2000). *Manual for the Patterns of Adaptive Learning Scales*. The University of Michigan. http://websites.umich.edu/~pals/PALS%202000_V13Word97.pdf

Nickow, A., Oreaopoulos, P., & Quan, V. (2020). *The impressive effects of tutoring on preK-12 learning: A systematic review and meta-analysis of the experimental evidence*. National Bureau of Economic Research. https://www.nber.org/papers/w27476

NWEA. (2020). *2020 NWEA MAP Growth normative data overview*. https://teach.mapnwea.org/impl/MAPGrowthNormativeDataOverview.pdf

What Works Clearinghouse. (2022). *What Works Clearinghouse standards handbook, version 5.0*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5_0-0-508.pdf

# Appendix A. Detailed description of Air Tutors math tutoring

Air Tutors provides live online tutoring on foundational math skills to middle school students in small groups. The program relies on high-quality tutors to deliver instruction and focuses on student and family engagement. Air Tutors aims to improve student math performance and math enjoyment, to generate positive student–tutor relationships, and to increase the number of students who are ready for Algebra I by 9th grade. Air Tutors focuses on three primary components:

1. **Live online tutoring.** Tutors provide instruction on foundational math skills, including addition, subtraction, decimals, fractions, proportions, and basic geometry. Students are tutored in small groups, which are assigned in groups of three or four (but frequently become smaller due to students not attending sessions). The curriculum uses the i-Ready curriculum scope and sequence and question sets.
2. **High-quality tutors.** Air Tutors recruits tutors who hold at least a bachelor's degree and three years of tutoring or teaching experience, with a preference for tutors who hold at least a master's or Ph.D. with five or more years of tutoring or teaching experience. Air Tutors also specifically seeks tutors who have engaging online personalities to best connect with students and who are relatively tech savvy to use the online platform efficiently.
3. **Student and family engagement.** Air Tutors program managers and tutors communicate with students and parents via text messages, emails, and phone calls. Tutors also determine when students might need a break from tutoring and implement noninstructional days (called Ubuntu sessions) to foster a positive culture with students.

The combination of tutors' engaging online personalities, small-group interaction with students, noninstructional days, family engagement, and text messaging communications with students was expected to lead to strong student–tutor relationships during the course of tutoring. These positive relationships, in turn, were expected to improve students' confidence and feelings of belonging in math in the short and long term. Tutors also provided personalized course content based on students' current math knowledge using a variety of instructional mediums, which was expected to lead to increased student math knowledge by the end of the semester. With this increased short-term math knowledge, students are expected to access more challenging math content in the future, which, in turn, is expected to lead to increased long-term math knowledge and more students who are prepared to take Algebra I in grade 9.

# Appendix B. Methods

**Experimental design.** This study used a randomized controlled trial (RCT) to analyze the impacts of Air Tutors on math knowledge outcomes, pre-post analysis to measure changes in student math confidence, and descriptive analysis (with no comparison group) to measure attendance, student–tutor relationships, and the context in which Air Tutors implemented the solution. Air Tutors implemented its program across three schools, including an online school (District 1), and an elementary school and middle school in a school district in Texas (District 2). Air Tutors invited 208 students to participate in its tutoring program across grades 4–6, including 64 students from District 1 and 144 students from District 2 (Table B.1). Eligible students included those who were available during tutoring times, were not receiving other additional services (District 1), did not have Individualized Education Programs (District 1 and the middle school in District 2), and were up to three grade levels below average on the Measures of Academic Progress (MAP) assessment (District 2).

**Table B.1. Number of treatment and control group students who were randomly assigned and who have outcome data, by site**

| Site | Randomly assigned | | | With outcome data | | |
|---|---|---|---|---|---|---|
| | Treatment | Control | Total | Treatment | Control | Total |
| All sites | 208 | 101 | 309 | 184 | 92 | 276 |
| District 1 online school | 64 | 38 | 102 | 52 | 33 | 85 |
| District 2 | 144 | 63 | 207 | 132 | 59 | 191 |
| District 2: Elementary school | 96 | 42 | 138 | 91 | 39 | 130 |
| District 2: Middle school | 48 | 21 | 69 | 41 | 20 | 61 |

Source: Program data.

Fourteen tutors provided tutoring in District 1, and 31 tutors provided tutoring in District 2. On average, students were offered 94 tutoring sessions (70 hours) across the 21-week program in District 1 and 55 tutoring sessions (42 hours) across the 15-week program in District 2. Sessions for students in District 1 spanned the fall 2021 and the spring 2022 semesters, while sessions for District 2 students took place only in spring 2022. Among invited students, 153 (74 percent) attended at least one session and 55 never attended a session. Reasons for never attending a session included scheduling conflicts and enrollment in a different school. Participation in Air Tutors was highest at the middle school in District 2 (88 percent), where students joined their online tutoring sessions from their school during the school day. Participation in District 1, where students attended tutoring sessions during their online school day, was similar to the overall average (73 percent). Participation at the elementary school in District 2, where tutoring sessions took place after school, was the lowest (67 percent).

**Measures.** We collected data from several sources before and after the study period to measure the impact of tutoring. From districts, we collected fall and spring scores on the MAP math assessment and student demographics. We collected survey data using the Patterns of Adaptive Learning Scale (PALS) and Math and Me to measure confidence in math and Copilot-Elevate to measure sense of belonging and student–tutor relationships (Adelson & McCoach, 2011; Midgley et al., 2000). Students participating in the study took the fall MAP test prior to the start of the study and the spring MAP test after the end of the tutoring program. MAP scores were standardized using grade-level national means and standard deviations provided by NWEA (2020). To assess the reliability of

survey measures, we calculated the Cronbach's alpha for each survey measure, and all measures exhibited values above the U.S. Department of Education's What Works Clearinghouse acceptable threshold of 0.6 (What Works Clearinghouse, 2022). The internal consistency (Cronbach's alpha) for math confidence and student–tutor relationships was high (greater than 0.8). The internal consistency for sense of belonging in math class, sense of belonging in tutoring sessions, feedback for growth, and student voice ranged from 0.6 to 0.7. We also collected attendance data for tutoring sessions from Air Tutors.

**Analysis.** The regression models for the math outcome (MAP) analysis all include the strata used in the randomization process, fall MAP scores, and race and ethnicity covariates. Because this is an RCT, impacts can be interpreted causally. We estimated findings for both participating students and all invited students. For invited students, we compared outcomes for all treatment group students to all control group students, regardless of whether students participated in the program (known as an intent-to-treat estimate). We estimated the impact on participating students by dividing the intent-to-treat estimate by the impact of randomization on participation using a two-stage equation (this is known as the treatment-on-the-treated estimate). The MAP analyses included 276 students (184 treatment group students, including 143 participating students and 92 control group students) who had spring and fall MAP scores (Table B.1). We also estimated Bayesian impact estimates using the BAyeSian Interpretation of Estimates (BASIE) tool, which implements Bayesian analysis in a standardized way using information from the What Works Clearinghouse database to inform a "prior distribution" that reflects relevant past research findings (Deke et al., 2022). These estimates systematically incorporate information from hundreds of other education studies on math outcomes, which inform the probability that the current study improved student outcomes. A baseline equivalence analysis revealed no significant differences in student characteristics before the program began (Table B.2).

The end-of-year survey analyses included 105 students participating in tutoring: 70 from District 2 who answered both the baseline and end-of-year surveys, and 35 from District 1 who answered both the baseline and end-of-year surveys. The interim survey analyses included 42 students from District 1 who responded to the baseline and interim surveys. The attendance analyses included the 153 students who participated in at least one session.

**Table B.2. Characteristics of treatment and control group students at baseline**

| Student characteristic | Overall | Treatment | Control | Difference | Standardized difference |
|---|---|---|---|---|---|
| Fall math MAP score (standardized) | -0.65 | -0.64 | -0.67 | 0.02 | 0.03 |
| Black | 21% | 20% | 22% | -2% | -0.06 |
| Hispanic or Latino | 67% | 68% | 64% | 4% | 0.10 |
| Number of students | 276 | 184 | 92 | 276 | 276 |

Source: District data.

Note: The fall math MAP score has been standardized so that a score at the national average would have a value of 0 and so the standard deviation is 1. Negative fall math MAP scores indicate scores lower than the national average. Standardized difference used Cox's index for binary variables and Hedge's g for continuous variables, following the What Works Clearinghouse (2022). We considered each characteristic to be balanced between treatment and control groups if the standardized difference was less than 0.25 standard deviations as stated in the *What Works Clearinghouse Standards Handbook, Version 5.0*.

* Significantly different from the control group at the .10 level, two-tailed test.

** Significantly different from the control group at the .05 level, two-tailed test.