

Social Programs That Work Review

## Evidence Summary for Saga Math Tutoring in Low-Income High Schools

### HIGHLIGHTS:

- **PROGRAM:** The Saga tutoring program provides math tutoring services to 9<sup>th</sup> and 10<sup>th</sup> graders in low-income high schools – 50 minutes per day, 5 days per week, for a full school year. Tutors are typically well-trained recent college graduates. Two closely-related versions of the program have been rigorously evaluated: Saga’s traditional program (with a student-tutor ratio of 2:1) and Saga Technology (with a student-tutor ratio of 4:1, supplemented by engagement with an education technology platform).
- **EVALUATION METHODS:** Three well-conducted randomized controlled trials (RCTs) evaluated Saga tutoring as implemented at scale in low-income public high schools in Chicago and/or New York City. Two of the studies evaluated Saga’s traditional program and the other evaluated the Saga Technology program.
- **KEY FINDINGS:** All three RCTs found that the program produced sizable, statistically significant effects on student math scores on the district tests at the end of the tutoring year. The effects represented approximately a 50-86% improvement over the annual gain in math otherwise expected of students in 9<sup>th</sup> or 10<sup>th</sup> grade. Two of the studies estimated the program’s effects at the end of 11<sup>th</sup> grade – one to two years after program completion – and found *sustained* sizable, statistically significant effects on district test scores. The effects represented approximately a 71% improvement over the annual gain in math otherwise expected of students in 11<sup>th</sup> grade.

[Disclosure: Arnold Ventures helped fund study 3, described in this summary.]

### I. Evidence rating: **TOP TIER**

The standard for Top Tier is:

*Programs shown in well-conducted RCTs, carried out in typical community settings, to produce sizable, sustained effects on important outcomes. Top Tier evidence includes a requirement for replication – i.e., the demonstration of such effects in two or more RCTs conducted in different implementation sites, or, alternatively, in one large multi-site RCT. Such evidence provides confidence that the program would produce important effects if implemented faithfully in settings and populations similar to those in the original studies.*

## **II. Description of the Program:**

Saga tutoring is an intensive, in-school math tutoring program for 9<sup>th</sup> or 10<sup>th</sup> grade students at all levels of math proficiency – not just struggling students. The program is delivered by Saga Education, a nonprofit organization, in low-income high schools.<sup>1</sup> Two closely-related versions of the program have been rigorously evaluated and found effective: Saga’s traditional program (evaluated in studies 1 and 2, below) and Saga Technology (evaluated in study 3, below).

Saga’s traditional program provides 9<sup>th</sup> or 10<sup>th</sup> graders with 50 minutes of in-school tutoring per day, 5 days per week, for one school year. The tutoring typically replaces an elective course (e.g., art or physical education), or a second hour of math (e.g., “double-dose” algebra). Tutors are well-educated individuals – often recent college graduates – who have strong math and interpersonal skills, and are willing to work for one year for a modest stipend as a public service. They receive training from Saga Education, and tutor two students at a time. The estimated program cost per student who receives tutoring is \$4,800.

Saga Technology is a lower-cost adaptation of Saga’s traditional model that incorporates an education technology component. Students in the program – 9<sup>th</sup> or 10<sup>th</sup> graders – alternate, on successive days, between working with an in-person tutor and individually engaging with an education technology platform (ALEKS). Tutor qualifications are similar to those in the traditional model, as are the program’s intensity and duration (50 minutes each school day for a full school year). Tutors work with four students at a time. The estimated program cost per student who receives the program is \$2,600.

[Saga Education’s website is linked here.](#)

## **III. Evidence of Effectiveness:**

This summary of the evidence is based on a systematic search of the literature to identify all well-conducted RCTs of Saga tutoring. Our search identified three such studies, summarized as follows.

### **STUDIES 1 AND 2 (Chicago Public Schools)**

#### **Study Design:**

Studies 1 and 2 evaluated Saga’s traditional model in 15 low-income public high schools in Chicago, with a combined sample of 5,343 incoming 9<sup>th</sup> and 10<sup>th</sup> graders in fall 2013 (study 1) and fall 2014 (study 2). The students were randomly assigned to a treatment group offered Saga tutoring versus a

---

<sup>1</sup> Saga was initially part of the Match charter school organization in Boston, and its tutoring program was named “Match tutoring.” It later became an independent organization – Saga Innovations, later renamed Saga Education – to focus on tutoring nationwide.

control group that received usual school services.<sup>2</sup> Over 95% of sample members were Black or Hispanic, and 87% were low income.

**Key Findings, Pooled Across the Two Studies:**

- **At the end of the tutoring year, the studies found a sizable effect on student math achievement as measured by district test scores.**
  - The effect size was 0.12 standard deviations and statistically significant ( $p < 0.01$ ). An effect of 0.12 represents approximately a 50% improvement over the annual gain in math otherwise expected in 9<sup>th</sup> or 10<sup>th</sup> grade.<sup>3</sup> This is the estimated effect for all students in the treatment group, regardless of whether or how long they actually received tutoring (i.e., it is the “intent-to-treat” effect).
  - However, only about 40% of students assigned to the treatment group received Saga tutoring due to scheduling problems and other issues. The study team therefore also estimated the program’s effect on treatment group students who received at least one tutoring session (i.e., the “treatment on treated” effect). This effect was 0.28 standard deviations and statistically significant ( $p < 0.01$ ). An effect of 0.28 represents approximately a 120% improvement over the annual gain in math otherwise expected in 9<sup>th</sup> or 10<sup>th</sup> grade.
  - These end-of-year effects were sizable and statistically significant in each of the two studies, analyzed separately.
- **The studies found *sustained* sizable effects on math achievement at the end of 11<sup>th</sup> grade – i.e., 1-2 years after program completion.**
  - The effect on district test scores, pooled across the two studies, was 0.10 for the full treatment group (intent-to-treat) and 0.23 for treatment group students who received tutoring (treatment on treated). Both effects were statistically significant ( $p < 0.01$ ). The 0.10 effect represents approximately a 71% improvement over the annual gain in math otherwise expected in 11<sup>th</sup> grade, and the 0.23 effect represents approximately a 160% improvement over this expected annual gain.<sup>4</sup>

---

<sup>2</sup> In study 1, students were randomly assigned to one of four treatment groups: (1) Saga tutoring, (2) Becoming a Man (BAM) – a program that provides elements of cognitive-behavioral therapy in weekly group sessions, (3) both Saga and BAM, or (4) neither program. In this summary, we focus on outcomes in the two Saga groups (as the treatment) versus the two non-Saga groups (as the control). The study found no discernible effects for the BAM group versus the “neither program” group on academic or behavioral outcomes.

<sup>3</sup> The average annual gain in math achievement for U.S. 9<sup>th</sup> and 10<sup>th</sup> grade students on six nationally normed tests is approximately 0.22-0.25 standard deviations (see Bloom, Hill, Black, and Lipsey, 2007). Saga’s impact of 0.12 represents approximately a 50% improvement over this annual gain.

<sup>4</sup> The average annual gain in math achievement for U.S. 11<sup>th</sup> grade students on six nationally normed tests is 0.14 standard deviations (see Bloom, Hill, Black, and Lipsey, 2007). Saga’s intent-to-treat effect of 0.10 represents approximately a 71% improvement over this annual gain, and its treatment-on-treated effect of 0.23 represents approximately a 160% improvement over this annual gain.

- These 11<sup>th</sup> grade effects were statistically significant in study 1 but not study 2, when the two studies were analyzed separately. Thus the evidence of sustained effects is strong but has not yet been definitively replicated.
- **The studies found short-term and sustained effects on math GPA but no discernible effects on most non-math outcomes.**
  - The studies found a statistically-significant effect on math GPA of 0.22 points during the tutoring year (the treatment group’s GPA was 2.02, versus 1.80 for the control group) and, in longer-term follow-up, a statistically-significant effect on 11<sup>th</sup> grade math GPA of 0.11 points (2.10 for the treatment group, versus 1.99 for the control group). These are effects for the full treatment group (i.e., intent-to-treat effects).
  - The studies found no discernible effects on most non-math outcomes, such as high school graduation rates, disciplinary suspensions, days absent, or a pre-specified index of such schooling outcomes.

**Discussion of Study Quality:**

Based on careful review, we believe these were both well-conducted RCTs. For example, the treatment and control groups were highly similar in their pre-program characteristics. Sample attrition rates for the outcomes described above ranged from 20% to 45% (depending on the outcome), and were nearly identical for the treatment versus control group. These rates constitute low attrition under the What Works Clearinghouse’s “cautious” standard ([WWC Procedures and Standards Handbook, Version 5.0, 2022](#)). The studies, in their main analyses, appropriately estimated Saga’s effects for all members of the treatment group, regardless of whether or how long they received program services (i.e., used an intent-to-treat analysis). The study analyses also appropriately adjusted for varying random assignment ratios across different schools and grades. The study appropriately [pre-registered](#) its primary outcomes, including student math achievement as measured by district test scores.

**STUDY 3 (Chicago and New York City)**

**Study Design:**

Study 3 evaluated the Saga Technology model in six low-income public high schools – three in Chicago and three in New York City. Within these schools, 2,065 incoming 9<sup>th</sup> graders in fall 2018 were randomly assigned to a treatment group offered the Saga Technology program versus a control group that received usual school services.<sup>5</sup> The sample was 24% Black and 57% Hispanic, and 91% were low income.

---

<sup>5</sup> The study also randomly assigned a second cohort of students entering 9<sup>th</sup> grade in fall 2019, but the COVID-19 pandemic disrupted program delivery and data collection for this cohort. Thus the study team specified, prior to analyzing outcomes for the full sample, that their primary analysis would focus on the first cohort only.

**Key Findings:**

- **At the end of 9<sup>th</sup> grade (the tutoring year), the study found a sizable effect on student math achievement as measured by district test scores – the study’s pre-specified primary outcome.**
  - The effect size was 0.19 standard deviations and statistically significant ( $p < 0.01$ ). An effect of 0.19 represents approximately an 86% improvement over the annual gain in math otherwise expected in 9<sup>th</sup> grade.<sup>6</sup> This is the estimated effect for all students in the treatment group, regardless of whether or how long they actually received tutoring (i.e., it is the “intent-to-treat” effect).
  - However, only 79% of students assigned to the treatment group received the Saga Technology program due to scheduling and other issues. The study team therefore also estimated the program’s effect on treatment group students who received at least one tutoring session (i.e., the “treatment on treated” effect). This effect was 0.23 standard deviations and statistically significant ( $p < 0.01$ ). An effect of 0.23 represents approximately a 105% improvement over the annual gain in math otherwise expected in 9<sup>th</sup> grade.
- **The study found sizable effects on other math-related outcomes during 9<sup>th</sup> grade and fall of 10<sup>th</sup> grade (before the COVID-19 disruption in spring of 10<sup>th</sup> grade).**
  - The study found a statistically-significant effect on math GPA of 0.20 points during the 9<sup>th</sup> grade year (the treatment group’s GPA was 2.13, versus 1.93 for the control group) and a statistically-significant effect on math GPA of 0.12 points during the fall of 10<sup>th</sup> grade (1.81 for the treatment group, versus 1.69 for the control group). These are effects for the full treatment group (i.e., intent-to-treat effects).
  - The study found a statistically-significant reduction of 4 percentage points in the proportion of math courses failed in 9<sup>th</sup> grade (treatment group members failed 16% of their math courses versus 20% for the control group); and a statistically-significant reduction of 5 percentage points in the proportion of math courses failed in the fall of 10<sup>th</sup> grade (21% for the treatment group, versus 26% for the control group). These are effects for the full treatment group (i.e., intent-to-treat effects).
- **The study found no discernible effects on non-math outcomes in the 9<sup>th</sup> grade year**, such as overall GPA, non-math GPA, non-math course failures, reading achievement scores, absences, and days suspended.

**Discussion of Study Quality:**

Based on careful review, we believe this was a well-conducted RCT. For example, the treatment and control groups were highly similar in their pre-program characteristics. Sample attrition rates for the

---

<sup>6</sup> The average annual gain in math achievement for U.S. 9<sup>th</sup> grade students on six nationally normed tests is 0.22 standard deviations (see Bloom, Hill, Black, and Lipsey, 2007). Saga’s impact of 0.19 represents approximately an 86% improvement over this annual gain.

math outcomes described above ranged from 11% to 19% (depending on the outcome), and were nearly identical for the treatment versus control group. These rates constitute low attrition under the What Works Clearinghouse’s “cautious” standard ([WWC Procedures and Standards Handbook](#), Version 5.0, 2022). The study, in its main analyses, appropriately estimated Saga Technology’s effects for all members of the treatment group, regardless of whether or how long they received program services (i.e., used an intent-to-treat analysis). The study analyses also appropriately adjusted for varying random assignment ratios across different schools and sub-populations. The study appropriately [pre-registered](#) its primary outcome, and closely adhered to its pre-registered analysis plan in estimating program effects.

#### **IV. References:**

##### **Studies 1 and 2:**

Jonathan Guryan, Jens Ludwig, Monica P. Bhatt, Philip J. Cook, Jonathan M. V. Davis, Kenneth Dodge, George Farkas, Roland G. Fryer Jr., Susan Mayer, Harold Pollack, Laurence Steinberg, and Greg Stoddard (2023). “Not Too Late: Improving Academic Outcomes among Adolescents,” *American Economic Review*, vol. 113, no. 3, pp. 738–765.

##### **Study 3:**

Monica P. Bhatt, Jonathan Guryan, Salman A. Khan, Michael LaForest-Tucker, and Bhavya Mishra (May 2024). “Can Technology Facilitate Scale? Evidence from a Randomized Evaluation of High-Dosage Tutoring,” *National Bureau of Economic Research (NBER) Working Papers*, no. 32510.