2022–23 Annual Social Purpose Report

The Math Agency is a Washington based Social Purpose Corporation, which means that we are committed to our social mission above all else. One of the requirements of a Social Purpose Corporation is an annual report on our progress towards this goal.

Our mission is to close educational gaps in math in public schools. During the 2022-23 academic year, we made progress on this mission on several fronts. This progress can be roughly divided into three categories: *developing capabilities to increase student growth*, *closing gaps with our partner schools,* and *developing our ability to scale*.

Developing capabilities to increase student growth

We continue to develop the tools, infrastructure, and capabilities that we believe are essential to our mission. During the past year we made notable progress on the following:

Measuring our ability to change individual student trajectories

A key part of our program is personalizing our coaching for each student. This year we progressed in our ability to *measure* the impact of specific coaching changes to *individual* students' growth trajectories. An example is shown below, where we can see a clear correlation between a coach change and an increase in a student's growth rate.
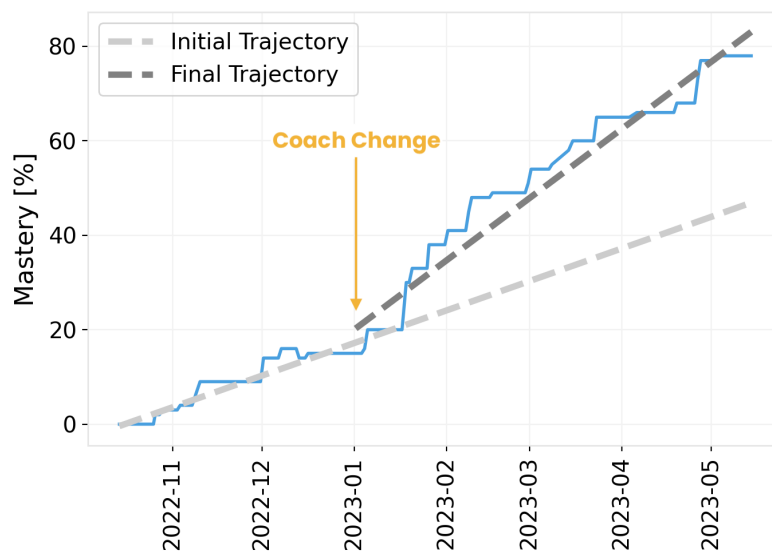

Figure 1. Plot of a student's skill mastery over time. A coach change is highlighted, and the dashed lines show piecewise linear fits to the mastery data. The change in growth rate was due to a combination of motivation (increased practice time) and efficiency (mastery per hour of practice).

We expect this ability will allow us to increase the impact of our program over time, especially for students who are showing lower growth than their peers. However, we are still early in the process of characterizing how effective and accurate these techniques are given the inherent noise in single-student data.

Measuring academic growth

During the 2022-23 academic year we gathered more evidence that our real-time measurement tools agree with standardized tests. We found further evidence that IXL's Assessment tool correlated strongly with the Measures of Academic Progress (MAP) assessment from NWEA and the Smarter Balanced Assessment (SBA). Data showing the relationship between MAP and IXL can be found in our [2022-23 Program Summary.](#)

Engaging families

This year we added text messaging as an additional tool for family communication. It had a significant impact on family response rates. Response rates increased from 15-20% for email-only communication to 20-40% after integrating text messaging.

As in previous years, we observed numerous cases of parent involvement driving student motivation. Generally this took the format of parents reminding or encouraging students to practice at home.

Closing gaps at our partner schools

The purpose of developing our technical capabilities is to have a positive impact on student learning. This year we continued to see strong evidence that this is indeed the case. We measured large increases in student growth via **four** separate measurements: mastery growth on Khan Academy, assessed skill levels on IXL, academic growth indices from the MAP, and via increases in the percentage of students meeting standards on the SBA.

Similar to our results from the 2021-22 year, we saw a *doubling* of academic growth rates for students enrolled in our program. On average, students entered having learned 0.7 grades/year of math, and they learned 1.4 grades/year while enrolled in our program (for programs that ran the full year). Additional details can be found in the 2022-23 program summary mentioned above.

One powerful feature of the MAP standardized assessment is that it provides direct Fall-Spring student growth metrics. Below, we show the MAP growth index for all 6 cohorts of students in our 2022-23 programs. The TOPS program was not a full-year program; it started in late January. We can see that in all of our full-year programs, students grew 130-220% more than students with similar grade levels and starting scores.
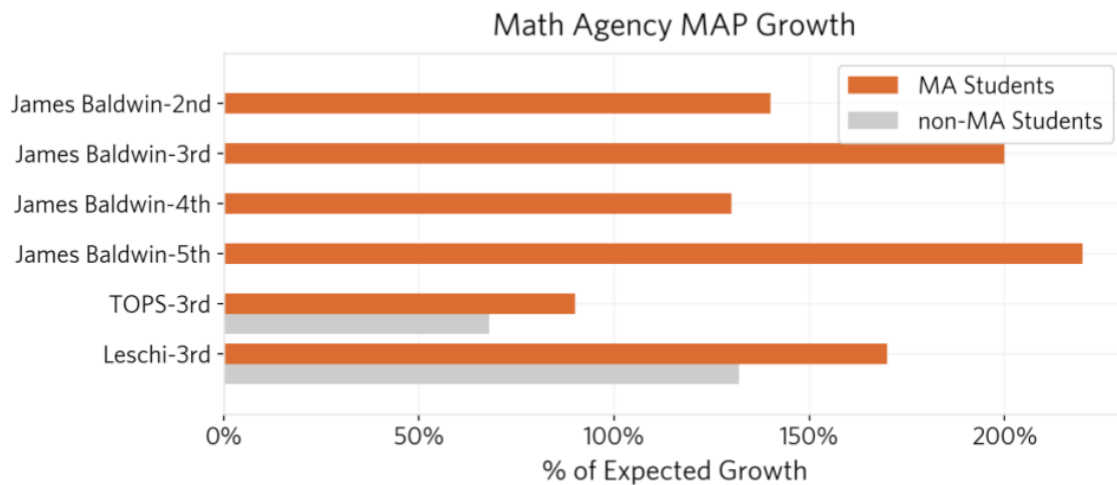

Figure 2. MAP growth index for our six student cohorts. We show data for non-enrolled students when available. On average, students enrolled in full-year programs (TOPS was a partial-year program) demonstrated 170% of expected growth.

While the SBA doesn't have specific growth metrics, it has the advantages of covering every school in the state and being publicly available via the Washington State Report Card website and data.wa.gov. The data is aggregated at the school-grade level, which means that it includes data from students in our program **and** students not enrolled in our program. One natural analysis is to look at the changes in the percent of students meeting standards as cohorts progress. For example, as 3rd graders become 4th graders, are more or less of them meeting standards? The results of our last two years of full-length cohorts are shown below in Figure 3.

On average, the six full-grade cohorts where we have a significant amount of student data have seen the amount of students meeting standards increase by over 26 percentage points. As can be seen in the table, all of the cohorts rank near the top of the state in terms of growth.

3

Cohort	Year	Met Standard	Prev. Met Standard	Change	WA Growth Rank
James Baldwin-4th*	2021-22	44.4%	16.1%	+28%	#13
James Baldwin-5th	2021-22	61.3%	16.1%	+45%	#2
Leschi-3rd*	2022-23	53.3%	40.4%	+13%	#143
James Baldwin-3rd*	2022-23	50%	23.3%	+27%	#14
James Baldwin-4th	2022-23	51.7%	23.3%	+28%	#8
James Baldwin-5th	2022-23	60.0%	44.4%	+16%	#27

*Figure 3. State assessment data (via the SBA) in math for different Math Agency cohorts. The change of a 2022–23 4th grade cohort would be calculated by looking at the difference of the 4th grade scores in 2022–23 relative to their 3rd grade scores in 2021–22. Note that James Baldwin was previously named Northgate. The Growth Rank column shows how each cohort ranks against similar (over 1k in total) cohorts in Washington state. *There isn't previous SBA data for these cohorts, and so we compare their scores to that of the same grade from the year before.*

To further put the SBA data into context, it is interesting to ask the following question: if we randomly chose six cohorts from the Washington state data, what are the odds that they would have a greater average SBA increase than our six? **The answer is ~1 in a million.** Of course, this isn't a completely perfect comparison, as we have a lot of overlap between schools and teachers in our data across the two years, but it does highlight the magnitude of the academic growth.

As always, there are important caveats to keep in mind regarding our growth data:
- Sample sizes: our sample sizes are growing, but still relatively small, and so we expect a material amount of measurement noise in our results.
- Causality: student growth is influenced by many things, including school-wide factors (building leadership, family engagement, after-school programs, etc.) and classroom factors (teachers and peer-related effects). We know that many of the students we work with benefit from great educators, and positive results are a result of the entire community.

Challenges to **fully** closing educational gaps

Our current goal is to get 80% or more of students at our partner schools to meet the basic standards in math. This is equivalent to the highest performing public schools in Washington today, and is our working definition of "closing educational gaps". We think this is possible based on two independent measures of impact:

-   Student-level growth: by analyzing [student growth trajectories](#), we can see that students enrolled in our program are on track to be at grade level within 2 years.
-   Public assessment data: we can see from Figure 3 that our partner schools have averaged over a 26% increase per year of students meeting standards. That suggests that 2 years of programming should correspond to roughly a 50% increase: enough to get the vast majority of schools to 80%.

However, to get 80% of students meeting standards requires us to be working with almost **all** of the at-risk students at our partner schools. This was challenging in 2022-23 for two reasons:

-   Coach capacity: at some of our partner schools, we did not have enough coaches to support all of the at-risk students in our targeted grades.
-   Student enrollment: at schools where we ran after-school programs, we were not able to get all of the at-risk students to enroll.

We are actively working to improve in both aspects in the upcoming year. In addition, we will be testing a variety of methods to increase our impact per student, such as increasing our dosage and improving how we choose content for students.

The ability to scale

It is very encouraging to see that we have not only maintained, but increased our per-student impact over the last two years: from +0.6 grades/year of extra growth in 2021-22 to +0.7 grades/year in 2022-23. That naturally leads to questions of scale: what are the barriers to helping *more students* with our program?

Right now we have two primary concerns: financial sustainability and coach supply. We made some progress on both over the course of the 2022-23 academic year.

Financial sustainability

Assuming we continue with our model of experienced site-directors paid at market rates, and assuming that they can serve 2-3 schools each, we expect that we'll need to charge between $1.5k and $2k per student (per year) to become financially sustainable. This

raises two questions. The first is around cost-effectiveness: i.e is it a good deal? The second is around market readiness: even if it is a good deal, are schools willing to pay for it?

Cost effectiveness

We find that schools and districts generally don't have strong historical evidence of effective gap-closing programs, which tends to render discussions of cost-effectiveness moot. However, there is a small amount of academic research on the subject. For example, a recent paper by Kraft and Falken explicitly calculates the cost-effectiveness of various tutoring programs. Their cost-effectiveness ratio (CER) is the increase in average per-student growth measured in standard deviations (SD), per thousand dollars spent. They find that CERs of published work range from 0.03 to 0.11 SD/$1k.

While we don't have direct effect size studies of our program, we do have several measurements of our impact on increased academic growth. We can use these to estimate our cost-effectiveness. For example, during the 2022-23 school year, we saw increased academic growth of +0.7 grades/year. Using the conversions of Lipsey et al, this corresponds to 0.3-0.4 SD, depending on the grade of the student. If we assume a program cost of $2k per student, this leads to a CER of between 0.15 to 0.2 SD/$1k, which is very encouraging[1]. This suggests that our program is very cost-effective. In other words, it **is** a good deal.

Market readiness

We are still in "learning mode" with regards to schools' and districts' willingness to pay for tutoring/coaching, even when it is cost-effective. So far a large part of our learning is that many schools and districts are not willing:) However, there are some encouraging recent developments. In particular, there has been a recent push for districts to use Outcome Based Contracts (OBCs) to align incentives around student learning. Some of the districts involved have published their associated rate cards, which provides a helpful benchmark on at least *some* districts' willingness to pay. For example, the 2023 Rate Cards show total payment ranges of $1.4k-$2.7k per student. This suggests that there is at least some market readiness for programs at our price point.

---

[1] The effect sizes used in Kraft and Falken are measured via randomly controlled trials (RCT), which tend to see smaller effect sizes than the less controlled methodology we are using in our effect size estimates. In this sense our measurements are likely overestimating our impact to some degree.

Coach supply

Even if our program becomes financially sustainable, it isn't clear to what extent our partner schools can find enough coaches to support students. Prior to the 2022-23 year, we did not take an active role in providing coaches for our program: we relied on our partner schools to provide the coaches: either via paid positions or with volunteers. In contrast, this year we worked closely with our school partners to develop a supply of well-trained volunteer coaches. These coaches come from a diverse set of backgrounds. They include family members at the school, high schoolers aiming to fulfill service learning hours, and retirees looking to give back to the community. We are gradually getting better at training and supporting them, and also at helping to refer them to our partner schools. We find that between 40-50% of the potential volunteers that reach out to us eventually end up coaching students at one of our partner schools.

We are still exploring the limits on the volunteer coach pool for our partner schools. Based on our observations, the coaches have been high quality, especially as observed by student growth metrics. It is important that we make progress understanding the *size* of this labor pool, as it has the potential to quickly become a bottleneck to serving more students. However, early signs are very encouraging, as can be seen below.
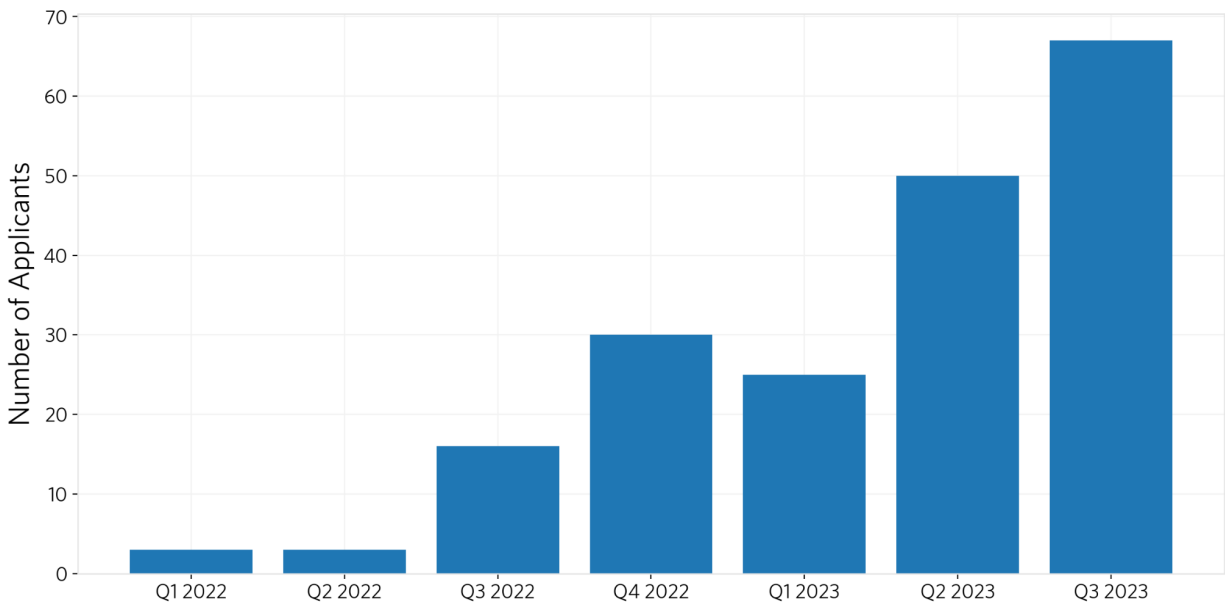

Figure 4. New coaching applicants per quarter over time.

Summary

During the 2022-23 academic year, The Math Agency made significant progress both on building infrastructure to increase our future impact and on demonstrating immediate progress on closing educational gaps in math for students in public schools. In particular, we believe that our high effectiveness, along with a very high *cost-effectiveness*, places us in a strong position for continued success in our mission.