Teaching Teachers to Use Computer Assisted Learning for Facilitating Personalized Learning at Scale:

Experimental and Non-Experimental Evidence

Philip Oreopoulos University of Toronto
Chloe Gibbs University of Notre Dame
Michael Jensen University of Notre Dame, LEO
Joseph Price Brigham Young University

Motivation

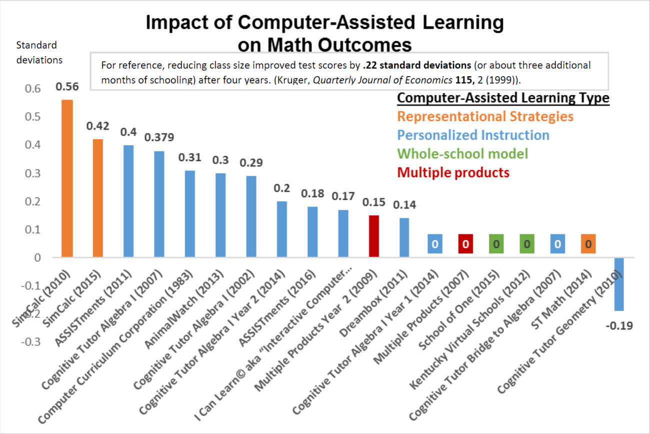
- Fundamental challenge of education is skill heterogeneity
- COVID-19 pandemic has exacerbated problem (e.g. Goldhaber et al., 2022, Engzell et al. 2021, Hanuchek, 2020)
- Pedagogical solution is personalized instruction but this approach often expensive

Supplemental Tutoring

- One-on-one instructional programming by teachers, paraprofessionals, volunteers, or parents, allows students to progress at their own pace and receive immediate feedback and attention
- Tutoring is expensive and difficult to scale, but overall benefits are impressive in a variety of settings
- A review of 96 randomized trials found consistent and substantial positive impacts on learning outcomes, with an average effects size of 0.37 standard deviation increase in test scores (usually english or math)

Using CAL to scale personalized learning

- Computer Assisted Technology (CAL) offers potential for simulating tutoring experience
- Example: Khan Academy
  - Roadmap of incremental short math videos and exercises to follow for Grades 3 to 12
  - Can progress at own pace
  - Receive immediate feedback and help
  - Mastery approach - keep trying until successfully completing questions
  - Roadmap can be customized for each student
  - Teacher, parents, tutors can observe student progress and respond if student is stuck



Small scale pilot programs, mostly in developed countries suggest CAL could be effective

Challenges with scaling CAL

- Students not motivated to use it - low take-up outside of school (e.g. Beg et al. 2022))
- Incorporating technology requires a change to the education production function by those delivering education
- Teachers and parents not familiar with it, concerned about additional effort and time costs to learn it, skeptical to adopt compared to using previous year's curriculum, too busy
- Sometimes limited computer access, in school and at home

Proposed Solution: Teach Teachers How to Use CAL Effectively

- Motivate, educate, and scaffold teachers for using CAL effectively as part of their curriculum.
- Key ingredient is assigning them a coach ('Khollaborator') to help train and support them throughout the year
- "Khollaborators" meet with teachers weekly to set goals, troubleshoot issues, and discuss best practices
- They also help monitor and interpret data, identify to teachers students that need attention
- Khollaborators help construct a roadmap of KA activities and videos for students to follow incrementally. Students work at their own pace
- Program is flexible: khollaborator works with teacher to modify roadmap, use assignments, and/or customize roadmap for each student

Road Map for Rest of this Presentation

- **Experiment 1: one week student-level 'proof of concept' experiment in Nashville TN**
  - test scores improve from KA practice, but only in classrooms with sufficienty high average practice times
- **Experiment 2: one school year teacher-level field experiment in Arlington TX**
  - test scores improve for elementary school students but no or neg. impact for middle school students - main difference appears to be less time to practice in middle school
- **Quasi-experimental evidence from 'as good as random' assignment to Arlington teachers**
  - test scores improve for students in classrooms where teacher facilitates sufficiently long average practice times
- **Conclusion and Next Steps: program shows promise but requires facilitating at least an hour a week of practice**

Nashville Experiment Setting

- In March 2022, a trial experiment to measure Khan Academy's ability to teach was run in 7 Metro Nashville Public Middle Schools (Schools of Innovation) with about 50 math teachers.
- 'Proof of concept' goal of demonstrating 'week in the life of KWiK' could improve math performance
- 3183 potential participants in grades 6-8 were randomly assigned 1 of 2 grade-specific topics to learn over the course of a week.
- During class, students took a pre-test on both topics, and they were then given 1 hour of Khan Academy activities and videos on their topic to complete that week.
- 1 week later, teachers gave class time to wrap up KA practice and students took a similar post test.

Nashville Experiment Setting

METRO NASHVILLE PUBLIC SCHOOLS

WILSON SHEEHAN LAB FOR ECONOMIC OPPORTUNITIES

This activity is designed to help reinforce and improve your ability to understand some key math concepts, and to help get you ready for your upcoming TCAP Achievement Tests.

1) Take a 6-question quiz to test your initial knowledge of the material.

- If you don't know a question, type **"Skip"** and move on.
- You won't need a calculator, but you may want scratch paper to work on.

2) Use the given username and password to log into Khan Academy and complete a series of videos and exercises.

Happy learning!

Next

Nashville Experiment Setting

Solve for h.

17 + 4h + 2 = 1 - 5h

Enter your answer as an integer.

Next

Grade 8, Topic A: Solving Equations
(Topics selected by *MNPS SOI's* Lead Math Instructional Specialist)

Nashville Experiment Setting

Now let's practice on Khan Academy!

1) Put on your headphones

2) Click HERE to sign into Khan Academy. Below are a specific username and password that you should use to sign in (you may need to log out of another account first)
Username:
Password:

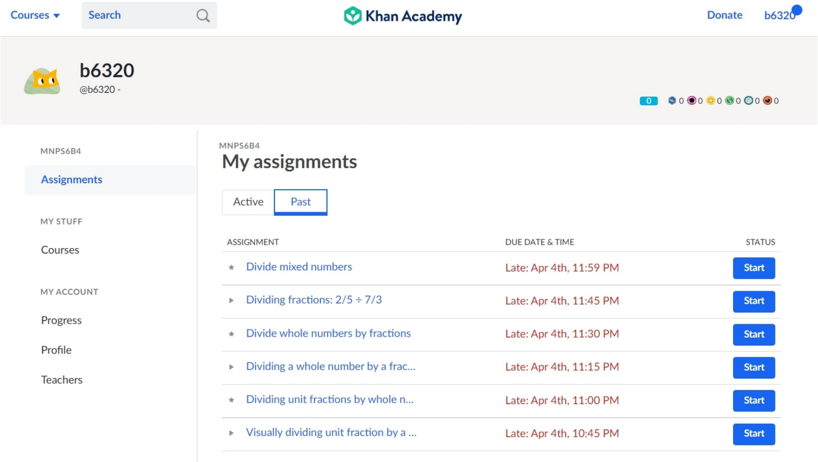
3) Start at the top and proceed sequentially until you complete the last exercise. Use paper and pen or pencil if you need to.

4) Don't feel frustrated if you have to try an exercise more than once before levelling up - learning from mistakes is how we improve.

5) Ask your teacher for help if you get stuck on an exercise or need a reminder of the instructions.

If you need more time, you can use free time this week to finish your practice.

Nashville Experiment Setting



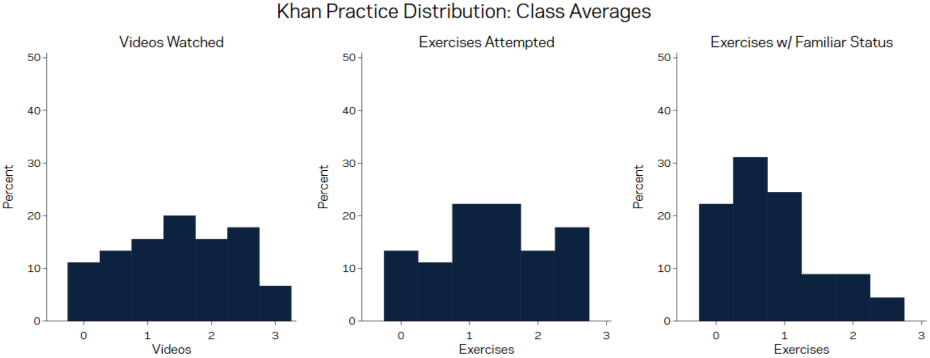
Grade 6, Topic B: Dividing Fractions

MNPS: Progress through Practice

	N	Exercise 1 Watched	Exercise 1 Attempted	Exercise 1 Familiar	Exercise 2 Watched	Exercise 2 Attempted	Exercise 2 Familiar	Exercise 3 Watched	Exercise 3 Attempted	Exercise 3 Familiar
Grade 6										
Topic 1	191	0.75	0.72	0.66	0.68	0.65	0.40	0.54	0.40	0.22
Topic 0	184	0.59	0.56	0.41	0.49	0.43	0.27	0.69	0.49	0.22
Grade 7										
Topic 1	222	0.66	0.67	0.45	0.52	0.37	0.14	0.36	0.30	0.22
Topic 0	204	0.72	0.68	0.50	0.54	0.50	0.19	0.41	0.33	0.15
Grade 8										
Topic 1	172	0.79	0.64	0.39	0.58	0.53	0.32	0.51	0.53	0.38
Topic 0	157	0.80	0.75	0.54	0.64	0.69	0.55	0.62	0.48	0.32
Total	1130	0.71	0.67	0.49	0.57	0.52	0.30	0.51	0.41	0.25
Average Classmate # Familiar										
Bottom 5%	75	0.53	0.41	0.20	0.28	0.25	0.09	0.19	0.15	0.03
Top 5%	79	0.95	0.92	0.78	0.81	0.78	0.71	0.75	0.68	0.61

Notes: Using only students with both pre and posttests. Familiar status signifies score above 70%. The activities were presented in order, alternating between videos and exercises to progress through practice. Videos were between 2-7 minutes and exercises consisted of 4 or 7 exercises.

MNPS: Various practice levels between classes



Notes: Distribution of 45 classes shown. Each class represents 1 data point (classes are not weighted by size). Familiar status represents a score above 70%.

MNPS: Model

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 X_{it} + \epsilon_{it}$$

- 2 observations for each student, one for their how they performed on treated topic and one for control.
- Y_{it} is the standardized posttest score for student i on topic t
- X_{it} is the standardized pretest score for student i on topic t
- T_{it} is the treatment status of topic t for student i
- ϵ_{it} is the error term, clustered at the student level

Note: We produce an alternative model that includes a student fixed effect. This does not significantly change coefficients or significance.

MNPS: Results

	(1)	(2)	(3)	(4)
		By Grade		
	Full Sample	Grade 6	Grade 7	Grade 8
Treatment	0.22***	0.29***	0.09*	0.30***
	(0.04)	(0.06)	(0.06)	(0.07)
N	2260	750	852	658

	(5)	(6)	(7)	(8)
	By Average Class Practice			
	<5 Mins.	5-25 mins.	25-50 mins.	50-100 mins.
Treatment	0.06	0.12	0.25***	0.28***
	(0.23)	(0.07)	(0.05)	(0.07)
N	50	480	1156	566

	(9)	(10)	(11)	(12)
	By Average Class Level-ups			
	<0.4 skills	0.4-0.8 skills	0.8-1.2 skills	>1.2 skills
Treatment	0.07	0.09	0.31***	0.36***
	(0.08)	(0.06)	(0.07)	(0.07)
N	386	622	586	666

	(13)	(14)	(15)
	By # of Activities Mastered		
	Exercise 1	Exercises 1-2	Exercises 1-3
Treatment	0.38***	0.43***	0.52***
	(0.05)	(0.07)	(0.08)
N	1110	640	436

Notes: Each student represents 2 observations. Scores on practiced and unpracticed topics are used for treatment and control, respectively. OLS estimations on the standardized posttest scores, with SEs clustered at the student level. All regressions control for standardized pretest scores. Class averages calculated leaving out the individual student.

With Student FEs | Lower bound | Upper bound

Takeaways from Nashville Experiment

- Using KA for reinforcing increasingly challenging math concepts significantly improves performance
- KWiK fidelity is critical: teachers need to facilitate sufficient practice time for students to master material (try again until levelling up)
- Effect sizes are impressively large when this happens (similar to high dosage tutoring)
- Caveat 1: short time horizon between practice and test
- Caveat 2: Counterfactual is not practicing, either because a) student was assigned to a different topic; b) time intended for practice used inefficiently (either due to teacher or student); c) teacher chose to use time differently

Arlington Experiment Setting

- At the end of the 2020-21 school year, we contacted all grade 3-8 math teachers in Arlington Independent School District about interest in using a CAL program in their class the following year
- 312 teachers expressed interest. They were grouped into 180 school/grade units to limit spillover between co-teachers
- After randomization (stratified by grade), 160 control teachers were selected to receive the program during the 2022-23 school year and contact was ceased

AISD: Sample Size

- From our original 312 teachers, AISD reported 253 with rostered students during the 2021-22 school year.
- 234 of these taught a math class
- 8 teachers whose students are missing 2022 STAAR scores (below grade 3).
- We drop remaining online teachers (2 teachers).
- **Final sample: 224 teachers; 10,949 students**

Note: We dropped students who spent time in both a treatment and control class, and assigned students who were in ≥ 2 treatment or control classes to the class they were in the longest.

AISD: Onboarding & Training

- Treatment teachers were given training on KWiK during district professional development prior to the start of school
- 33 khollaborators were hired to provide support to teachers throughout the year.
- Khollaborators were trained to provide technical support, as well as keep teachers accountable towards goals
- Khollaborator/teacher meetings happened virtually; weekly to start the year, moved to monthly or less as teachers adjusted

Teacher completed onboarding survey	92.9%
Teacher attended professional development*	78.6%
Teacher attended first khollaborator meeting	90.2%
Teacher scheduled follow-up meetings**	76.8%

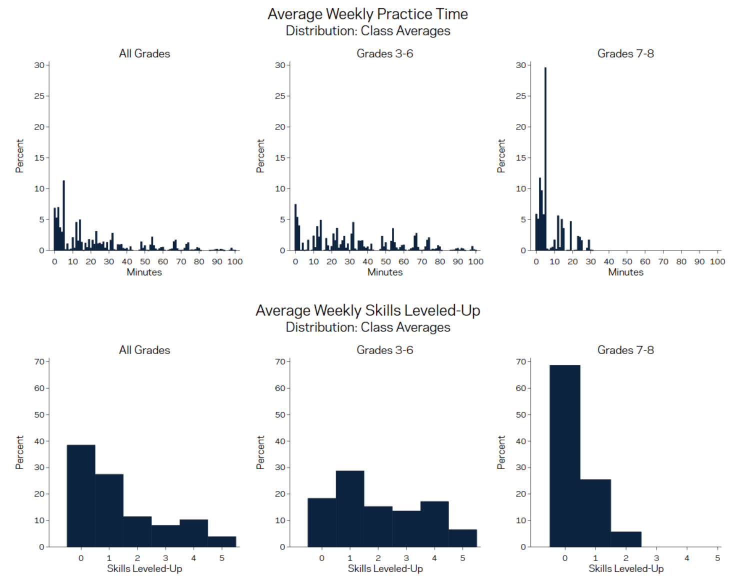
N: 112 treatment teachers. *Attended Live. Teachers also had option to watch recorded PD afterward. **Had follow-up meeting scheduled with khollaborator before Oct 7th, 2021

AISD: Grade Divisions

From the start, there was evidence supporting treating grades 3-6 (elementary) and 7-8 (middle) differently:

- Elementary classrooms blocked for 70 minutes per day for math, but teachers had flexibility to adjust since students remained in the same classroom for multiple subjects.
- Elementary classrooms had additional enrichment time throughout the week that teachers could use for any subject
- Middle school classrooms had 50 minutes per day with their math students
- Grades 3-6 average students per teacher: 43.7; Grades 7-8: 88.3
- Many teachers expressed disinterest in assigning Khan Academy as additional homework, especially given lack of tech access in some homes

AISD: Participation Distribution by Class



Control Group

AISD: Descriptive Stats, Balance

	N	Control	Treatment	Diff
Age	156	10.05	10.09	0.034
Race: White	156	0.51	0.53	0.022
Race: Black	156	0.29	0.26	-0.030
Ethnicity: Hispanic	156	0.52	0.47	-0.051
Race: Native American	156	0.10	0.10	-0.005
Female	156	0.50	0.46	-0.035***
Days Missed	156	20.53	20.19	-0.336
ESL	156	0.23	0.23	0.004
Special Ed	156	0.13	0.16	0.034**
Free lunch Eligible	156	0.50	0.47	-0.023
Grade 3	156	0.25	0.24	-0.013
Grade 4	156	0.17	0.21	0.036
Grade 5	156	0.20	0.17	-0.029
Grade 6	156	0.23	0.24	0.012
Grade 7	156	0.07	0.08	0.004
Grade 8	156	0.07	0.07	-0.009

Notes: Student characteristics, averaged among grade-school units. Represents 10,979 students among 224 teachers.

Balance: Teacher Survey | Stats by grade: Teacher Survey | Stats by grade: Treatment Onboarding Survey

AISD: Model

$$Y_{igs} = \beta_0 + \beta_1 T_{igs} + \beta_2 X_{igs} + \gamma_g + \epsilon_{igs}$$

- Y_{igs} is the standardized Math STAAR 2022 score for student i in grade g and school s
- T_{igs} is the treatment status of student i in grade g and school s
- X_{igs} is a matrix of person level controls for student i in grade g and school s
- γ_g is a grade level fixed effect, and ϵ_{igs} is the error term

AISD: Main Results

	I No Controls	II Grade FEs	III Grade FEs w/ Controls
All Grades	0.036	0.044	0.025
	(0.092)	(0.083)	(0.076)
N	10,979	10,979	10,979
Grades 3-6	0.171**	0.172**	0.122**
	(0.069)	(0.070)	(0.058)
N	7,234	7,234	7,234
Grades 7-8	-0.201	-0.202	-0.173
	(0.206)	(0.194)	(0.202)
N	3,745	3,745	3,745

Notes: OLS regressions of standardized 2022 Math STAAR scores on treatment. Standard errors clusted at the grade/school level. Controls for III include: age, sex, race, ethnicity, days missed, english learner status, special ed status, free lunch eligibility.

By individual grade

AISD: Subgroup Analysis (Grades 3-6)

		Gender			Race			Ethnicity		
Specification	All	Female	Male	F-Stat	White	Black	F-Stat	Hispanic	Not	F-Stat
No Controls	0.171**	0.200***	0.139*	1.56	0.139*	0.121	0.04	0.065	0.244**	4.06**
	(0.069)	(0.072)	(0.075)		(0.075)	(0.085)		(0.047)	(0.103)	
	7234	3507	3727		3794	2015		3434	3800	
Grade FEs	0.172**	0.201***	0.140*	1.56	0.139*	0.140*	0.02	0.068	0.249**	4.23**
	(0.070)	(0.071)	(0.076)		(0.075)	(0.084)		(0.047)	(0.102)	
	7234	3507	3727		3794	2015		3434	3800	
Grade FEs w/ Controls	0.122**	0.158**	0.085	2.63	0.081	0.163**	0.53	0.098**	0.139*	0.83
	(0.058)	(0.061)	(0.063)		(0.061)	(0.082)		(0.049)	(0.077)	
	7234	3507	3727		3794	2015		3434	3800	

Notes: OLS regressions of standardized 2022 Math STAAR scores on treatment, broken out by subgroup. Grades 3-6 only. Standard errors clusted at the grade/school level. Controls for III include: age, sex, race, ethnicity, days missed, english learner status, special ed status, free lunch eligibility. F-Stat comparing treatment coefficient from both subgroups.

More subgroups: Spec-Ed and Economic Indicator

Takeaways from Arlington Experiment

- Full sample program effect on student math scores insignificant
- This overall result masks distinct differences in estimates between elementary and middle schools, corresponding to distinct differences in practice time
- In elementary schools where average practice time was more than 30 minutes each week, state math scores were about 20% standard deviations higher for treated classrooms, but in middle schools where average practice time was less than 10 minutes each week, tests scores among treated students were significantly lower
- We posit that, given the little time middle school teachers have to teach math in Arlington, and their avoidance of homework, there is insufficient time for students to adopt mastery approach when practicing on KA, and this results in a worse substitute than the current curriculum
- Same conclusion as Nashville: KWiK fidelity is critical: teachers need to facilitate sufficient practice time for students to master material (try again until levelling up)

Caveats from Arlington Experiment

- Caveat 1: Breakdown of results by elementary vs middle school not in original pre-analysis
- Caveat 2: Estimates are imprecise, with p-values around 0.05
- Caveat 3: Even among elementary schools, high variance in fidelity

Quasi-Experimental Analysis

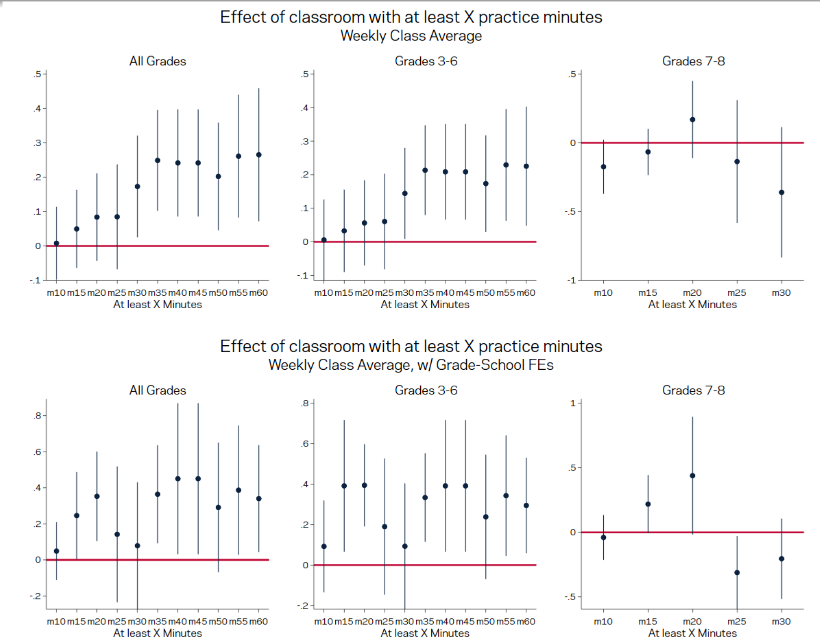
- Perhaps students ’as good as randomly assigned’ to teachers within schools and grades, at least conditional on past test scores
- Similar to teacher value added models, except we examine the relationship between students assigned to teachers that facilitated more practice time on KA than others
- Past research suggests teacher value added credibly identified after conditioning on past scores (estimates similar to cases with student RA) (Chetty et al. 2014, 2017), Koedel et al. 2015, Kane et al. 2013)
- Caveat: are good teachers same ones who facilitate more KA practice?

AISD: Value Added Model

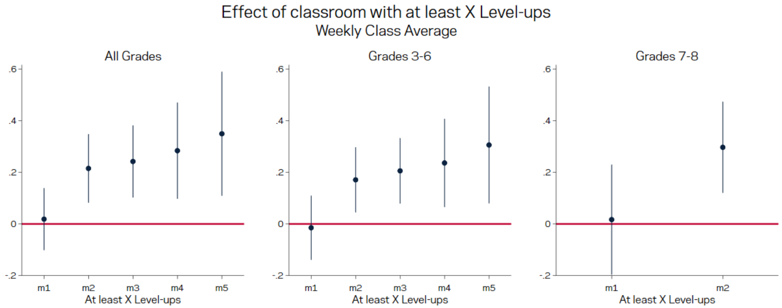
$$Y_{igs} = \beta_0 + \beta_1 P_{igs} + \beta_2 Y_{ig-1s} + \gamma_g s + \epsilon_{igs}$$

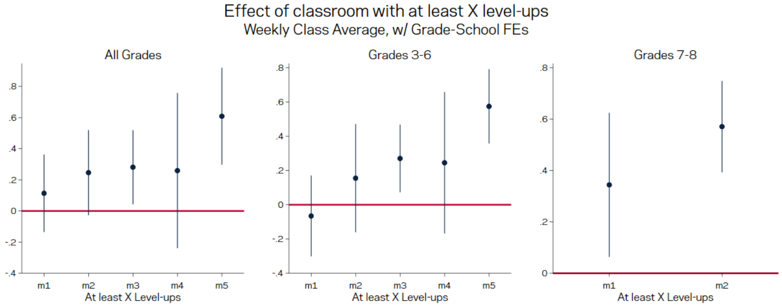
- Y_{igs} is the standardized Math STAAR 2022 score for student i in grade g and school s
- P_{igs} is an indicator for whether average classroom practice was at least P minutes (leaving student out) for student i in grade g and school s
- Y_{ig-1s} is previous year's test standardized test score (must drop Grade 3) for student i in grade g and school s
- $\gamma_g s$ is a schoolxgrade level fixed effect, and ϵ_{igs} is the error term

AISD: Quasi-Experiment Results



AISD: Quasi-Experimental Results





AISD: Onboarding Survey and Practice (Grades 3-6)

	Coeff	SE
Teacher attended PD session during district PD Day	21.20***	(6.44)
Teacher held first meeting with khollaborator	35.20***	(4.45)
Teacher held followup meeting with khollaborator prior to Oct 1st, 2021	26.90***	(8.37)
How many days per week will your students receive math instruction?	-2.48	(5.11)
About how long is each math class?	-0.10	(0.26)
How easy do you expect it would be for you to provide some time each week to work on Khan Academy DURING CLASS?		
I think I can provide some practice time once a week, for less than 30 minutes	-3.65	(17.29)
I think I can provide some practice time once a week, for more than 30 minutes	0.75	(10.19)
I think I can provide some practice time more than once a week, for less than 30	-1.48	(8.16)
I think I can provide some practice time more than once a week, for more than 30	9.18	(8.65)
How easy do you expect it would be for students to practice on Khan Academy AFTER SCHOOL for at least an hour a week?		
Easy	28.26*	(14.45)
I don't usually assign my students any homework	-20.20**	(7.85)
I only assign my students homework if they are unable to complete practice assig	-0.16	(7.36)
I usually assign homework each week	-2.94	(10.38)
Most students don't often complete their homework	2.53	(9.10)
I make homework a graded requirement (even just a participation grade)	10.26	(12.80)
Some students do not have access to use or borrow a computer or tablet for homew	-6.30	(6.78)
Students have access to a computer lab at school that they can use during lunch	-19.68***	(5.52)
Which homework guideline do you think would work for your students (and their parents) to try to follow?		
What practice time you don't finish during school, finish at home.	-2.15	(7.80)
Once a week, focus on practicing for at least an hour	-19.39*	(10.55)
Two times a week for at least 30 minutes each time	7.21	(7.77)
Try a little every day	4.14	(12.56)
Do at least an hour of practice, but see if you can do even more.	-15.16*	(8.05)
How easy do you expect it would be to communicate with parents about the program, asking for help to facilitate at least an hour of practice at home?		
Most parents read messages when I send them by email	7.33	(7.48)
Most parents read messages when I send them by text or use Remind	-2.02	(8.27)
Most parents read messages when I provide a handout for students to take home to	6.50	(8.39)
I have used the app Remind in the past to message parents	0.30	(9.31)
Most parents are too busy or too overwhelmed to help	0.43	(8.21)
Most parents do not want their children assigned homework	-7.50	(9.75)

Notes: OLS coefficients of a teacher's average weekly practice time on their answers to onboarding questions. SE clustered at grade student level, FEs for grade. N: 84 teachers (treatment, grade 3-6)

Grades 7-8 | Pre survey correlation w/ practice (grades 3-6 | Pre survey correlation w/ practice (grades 7-8)

Takeaways from Arlington Quasi-Experiment

- There exists a non-linear relationship in KA practice time among students with similar math backgrounds: only those averaging at least 30 minutes of practice each week or levelling up at least twice a week perform better by the end of the school year
- A similar relationship arises when looking at differences in performance by average classroom practice time and level ups
- Same conclusion as Nashville and Arlington Experiments: KWiK fidelity is critical: teachers need to facilitate sufficient practice time for students to master material (try again until levelling up)
- Caveat 1: assignment to classrooms may not be random
- Caveat 2: teachers who facilitate practice time could be better at in other ways

Conclusions so far

- Computer Assisted Learning platforms, such as Khan Academy, can simulate the tutoring experience, allowing students to progress incrementally at their own pace, receiving immediate feedback, and achieving mastery before advancing
- Training, supporting, and guiding teachers to adapt CAL can be an effective way to generating meaningful test score gains (at least for math)
- But critical that sufficient practice time be given to work on CAL in order for students to have time to repeat exercises, watch videos, and advance through material

Next Steps, Current Study

- Produce results for term performance
- Robustness check on quasi-experiment to see if teachers that facilitated practice also improved scores to previous year's students
- Survey and interview data to characterize teachers who facilitated practice vs ones that did not
- Year 2 data to look at longer run outcomes, and increase sample for quasi-experiment

Next Steps, New Studies

- KWiK may work better in districts that facilitate KA practice both in class and as homework (Toronto pilot underway)
- Create Professional Development materials to lower costs and scale
- Explore add-on virtual tutoring to help students that do not practice well or level up in class: link virtual tutoring to classroom homework
- Continue to identify reasons for teacher heterogeneity and explore ways for improving fidelity
- introduce ChatGPT

Appendix: MNPS Results with student FEs

	(1)	(2)	(3)	(4)
			By Grade	
	Full Sample	Grade 6	Grade 7	Grade 8
Treatment	0.22***	0.29***	0.08	0.30***
	(0.05)	(0.08)	(0.08)	(0.10)
N	2260	750	852	658

	(5)	(6)	(7)	(8)
	By Average Class Practice			
	<5 Mins.	5-25 mins.	25-50 mins.	50-100 mins.
Treatment	0.08	0.12	0.24***	0.27***
	(0.33)	(0.10)	(0.07)	(0.10)
N	50	480	1156	566

	(9)	(10)	(11)	(12)
	By Average Class Level-ups			
	<0.4 skills	0.4-0.8 skills	0.8-1.2 skills	>1.2 skills
Treatment	0.06	0.09	0.31***	0.35***
	(0.11)	(0.09)	(0.10)	(0.09)
N	386	622	586	666

	(13)	(14)	(15)
	By # of Activities Mastered		
	Exercise 1	Exercises 1-2	Exercises 1-3
Treatment	0.38***	0.43***	0.52***
	(0.07)	(0.10)	(0.12)
N	1110	640	436

Notes: Each student represents 2 observations. Scores on practiced and unpracticed topics are used for treatment and control, respectively. OLS estimations on the standardized posttest scores, with SEs clustered at the student level. All regressions control for standardized pretest scores and include student fixed effects. Class averages calculated leaving out the individual student.

BACK

Appendix: Lower Bound MNPS Results

We attempt to find a realistic lower bound for our main results by assigning any student who did not complete a post-test a score increase of 0 from their pre-test score. The average increase among those with both tests was 0.038 points.

	(1)	(2)	(3)	(4)
		By Grade		
	Full Sample	Grade 6	Grade 7	Grade 8
Treatment	0.14***	0.19***	0.08*	0.17***
	(0.02)	(0.04)	(0.04)	(0.04)
N	3612	1094	1236	1282

	(5)	(6)	(7)	(8)
	By Average Class Practice			
	<5 Mins.	5-25 mins.	25-50 mins.	50-100 mins.
Treatment	-0.03	0.07*	0.19***	0.23***
	(0.04)	(0.04)	(0.04)	(0.06)
N	368	1018	1486	728

	(9)	(10)	(11)	(12)
	By Average Class Level-ups			
	<0.4 skills	0.4-0.8 skills	0.8-1.2 skills	>1.2 skills
Treatment	0.03	0.07	0.22***	0.30***
	(0.03)	(0.05)	(0.06)	(0.06)
N	1084	898	812	818

	(13)	(14)	(15)
	By # of Activities Mastered		
	Exercise 1	Exercises 1-2	Exercises 1-3
Treatment	0.30***	0.33***	0.41***
	(0.04)	(0.06)	(0.07)
N	1432	792	512

Notes: Each student represents 2 observations. Scores on practiced and unpracticed topics are used for treatment and control, respectively. OLS estimations on the standardized posttest scores, with SEs clustered at the student level. All regressions control for standardized pretest scores. Class averages calculated leaving out the individual student.

BACK to Main Results Upper bound

Appendix: Upper Bound MNPS Results

For an upper bound on our results, we assign any missing a post test a score of 3 (out of 3) on their practiced topic, and a score equal to their pretest score for their unpracticed topic.

	(1)	(2)	(3)	(4)
			By Grade	
	Full Sample	Grade 6	Grade 7	Grade 8
Treatment	0.98***	1.03***	0.79***	1.11***
	(0.03)	(0.06)	(0.06)	(0.05)
N	3612	1094	1236	1282

	(5)	(6)	(7)	(8)
	By Average Class Practice			
	$<$5 Mins.	5-25 mins.	25-50 mins.	50-100 mins.
Treatment	1.93***	1.19***	0.72***	0.76***
	(0.09)	(0.06)	(0.05)	(0.07)
N	368	1018	1486	728

	(9)	(10)	(11)	(12)
	By Average Class Level-ups			
	$<$0.4 skills	0.4-0.8 skills	0.8-1.2 skills	$>$1.2 skills
Treatment	1.46***	0.71***	0.87***	0.76***
	(0.06)	(0.07)	(0.07)	(0.07)
N	1084	898	812	818

	(13)	(14)	(15)
	By # of Activities Mastered		
	Exercise 1	Exercises 1-2	Exercises 1-3
Treatment	0.77***	0.72***	0.70***
	(0.05)	(0.07)	(0.08)
N	1432	792	512

Notes: Each student represents 2 observations. Scores on practiced and unpracticed topics are used for treatment and control, respectively. OLS estimations on the standardized posttest scores, with SEs clustered at the student level. All regressions control for standardized pretest scores. Class averages calculated leaving out the individual student.

BACK to Main Results Lower bound

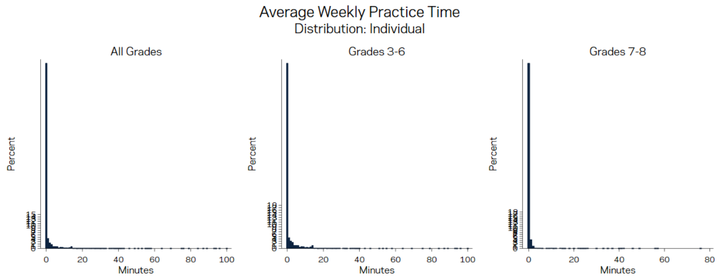
Appendix: MNPS Participation & Balance

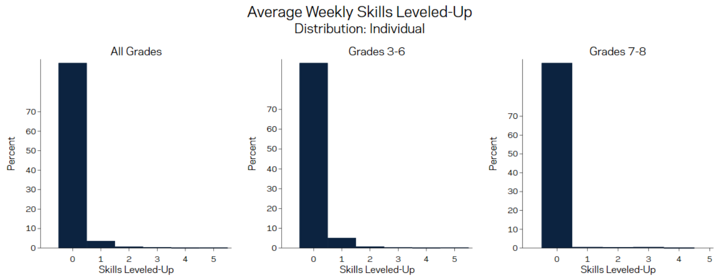
	N	Group A	Group B	Diff
Grade 6				
Took Pre-Test	996	0.55	0.55	-0.003
Took Post-Test	996	0.41	0.41	-0.007
Grade 7				
Took Pre-Test	1064	0.58	0.58	0.001
Took Post-Test	1064	0.47	0.41	-0.058***
Grade 8				
Took Pre-Test	1124	0.57	0.57	-0.001
Took Post-Test	1124	0.31	0.30	-0.014

Notes: The sample here differs from the previous table. We pre-randomized all grades 6-8 students, before knowing who would participate. This table includes all randomized students, while the previous shows all participating students.

BACK

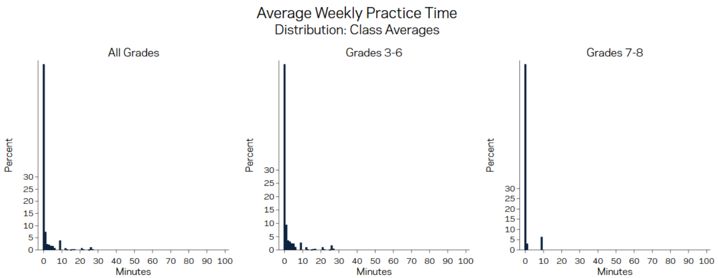
Appendix: Participation Distribution by Student (Control)

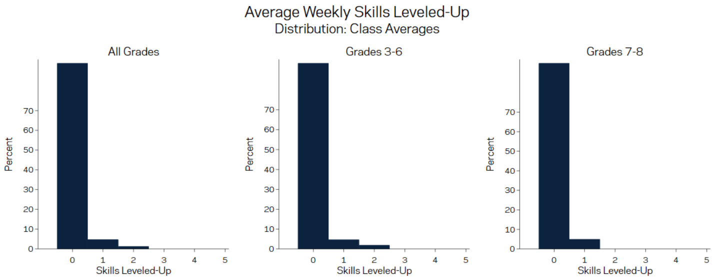




BACK

Appendix: Participation Distribution by Class (Control)





BACK

Appendix: AISD Treatment Onboarding Survey, Desc

	All Treatment	Grade 3-6	Grade 7-8
How many days per week will your students receive math instruction?	4.88	4.89	4.84
About how long is each math class?	64.23	69.18	48.60
How easy do you expect it would be for you to provide some time each week to work on Khan Academy DURING CLASS?			
I think I can provide some practice time once a week, for less than 30 minutes	0.19	0.11	0.44
I think I can provide some practice time once a week, for more than 30 minutes	0.29	0.25	0.40
I think I can provide some practice time more than once a week, for less than 30	0.49	0.57	0.24
I think I can provide some practice time more than once a week, for more than 30	0.25	0.30	0.08
How easy do you expect it would be for students to practice on Khan Academy AFTER SCHOOL for at least an hour a week?			
Easy	0.16	0.15	0.20
I don't usually assign my students any homework	0.25	0.25	0.24
I only assign my students homework if they are unable to complete practice assig	0.29	0.30	0.24
I usually assign homework each week	0.21	0.15	0.40
Most students don't often complete their homework	0.27	0.22	0.44
I make homework a graded requirement (even just a participation grade)	0.16	0.11	0.32
Some students do not have access to use or borrow a computer or tablet for homew	0.40	0.44	0.28
Students have access to a computer lab at school that they can use during lunch	0.04	0.04	0.04
Which homework guideline do you think would work for your students (and their parents) to try to follow?			
What practice time you don't finish during school, finish at home.	0.47	0.44	0.56
Once a week, focus on practicing for at least an hour	0.07	0.06	0.08
Two times a week for at least 30 minutes each time	0.25	0.25	0.24
Try a little every day	0.19	0.22	0.12
Do at least an hour of practice, but see if you can do even more.	0.02	0.03	0.00
How easy do you expect it would be to communicate with parents about the program, asking for help to facilitate at least an hour of practice at home?			
Most parents read messages when I send them by email	0.53	0.58	0.36
Most parents read messages when I send them by text or use Remind	0.48	0.52	0.36
Most parents read messages when I provide a handout for students to take home to	0.28	0.34	0.08
I have used the app Remind in the past to message parents	0.41	0.35	0.60
Most parents are too busy or too overwhelmed to help	0.28	0.25	0.36
Most parents do not want their children assigned homework	0.12	0.11	0.12
N	105	80	25

Notes: Responses from an onboarding survey taken by treatment teachers prior to the start of the school year. Teachers instructed to select all applicable answers. 105 of 112 treatment teachers responded.

BACK

Appendix: AISD Pre-randomization Survey, Desc

	All Treatment	Grade 3-6	Grade 7-8
Percent of Students with Access to Tech	84.62	84.74	84.20
Teaches multiple grades	0.03	0.03	0.02
Prior Khan Use			
Never Used KA	0.19	0.19	0.18
Use KA for Demonstration Only	0.45	0.45	0.44
Use KA for Optional Practice	0.23	0.22	0.28
Use KA for Required Practice	0.13	0.14	0.10
Student Independent work in School			
Students not Expected to Practice Math Indep. in School	0.00	0.00	0.00
Students Expected to Practice Math Indep. in School 1-30 mins weekly	0.21	0.23	0.12
Students Expected to Practice Math Indep. in School 31-60 mins weekly	0.28	0.25	0.38
Students Expected to Practice Math Indep. in School 61-90 mins weekly	0.24	0.22	0.30
Students Expected to Practice Math Indep. in School 90 plus mins weekly	0.28	0.30	0.20
Student Math practice after School			
Students not Expected to Practice Math After School	0.10	0.11	0.06
Students Expected to Practice Math After School 1-30mins weekly	0.50	0.54	0.36
Students Expected to Practice Math After School 31-60mins weekly	0.27	0.23	0.42
Students Expected to Practice Math After School 61-90mins weekly	0.09	0.09	0.12
Students Expected to Practice Math After School 90 plus mins weekly	0.03	0.03	0.04
Student KA use in School			
Students Never work on KA in school	0.35	0.34	0.38
Students work on KA less than 1x per week in School	0.36	0.32	0.50
Students work on KA about 1x per week in School	0.17	0.20	0.06
Students work on KA more than 1x per week in School	0.12	0.14	0.06
Student KA use after School			
Students Never work on KA after School	0.64	0.66	0.58
Students work on KA less than 1x per week after School	0.27	0.27	0.28
Students work on KA about 1x per week after School	0.06	0.05	0.10
Students work on KA more than 1x per week after School	0.03	0.02	0.04
Student non-KA CAT use in School			
Students never work on non-KA CAT in school	0.12	0.09	0.20
Students work on non-KA CAT less than 1x per week in School	0.11	0.07	0.22
Students work on non-KA CAT about 1x per week in School	0.15	0.14	0.18
Students work on non-KA CAT more than 1x per week in School	0.63	0.69	0.40
Student non-KA CAT use after School			
Students Never work on non-KA CAT after School	0.25	0.21	0.38
Students work on non-KA CAT less than 1x per week after School	0.26	0.27	0.24
Students work on non-KA CAT about 1x per week after School	0.19	0.21	0.12
Students work on non-KA CAT more than 1x per week after School	0.30	0.32	0.26
N	224	174	50

Notes: Averages for answers received on the pre-randomization survey, asking teachers about their prior experience with Khan Academy and classroom expectations for students.

BACK

Appendix: AISD Pre-randomization Teacher Survey, Balance

	All Grades N	Control	Treatment	Diff
Percent of Students with Access to Tech	156	84.87	83.52	-1.350
Teaches multiple grades	156	0.00	0.02	0.025*
Prior Khan Use				
Never Used KA	156	0.23	0.18	-0.047
Use KA for Demonstration Only	156	0.43	0.43	-0.002
Use KA for Optional Practice	156	0.19	0.24	0.046
Use KA for Required Practice	156	0.15	0.15	0.003
Student Independent work in School				
Students not Expected to Practice Math Indep. in School	156	0.00	0.00	0.000
Students Expected to Practice Math Indep. in School 1-30 mins weekly	156	0.19	0.24	0.049
Students Expected to Practice Math Indep. in School 31-60 mins weekly	156	0.27	0.26	-0.009
Students Expected to Practice Math Indep. in School 61-90 mins weekly	156	0.21	0.26	0.058
Students Expected to Practice Math Indep. in School 90 plus mins weekly	156	0.33	0.24	-0.099
Student Math practice after School				
Students not Expected to Practice Math After School	156	0.12	0.10	-0.021
Students Expected to Practice Math After School 1-30mins weekly	156	0.49	0.55	0.062
Students Expected to Practice Math After School 31-60mins weekly	156	0.27	0.22	-0.048
Students Expected to Practice Math After School 61-90mins weekly	156	0.08	0.09	0.003
Students Expected to Practice Math After School 90 plus mins weekly	156	0.03	0.03	0.004
Student KA use in School				
Students Never work on KA in school	156	0.36	0.31	-0.047
Students work on KA less than 1x per week in School	156	0.31	0.39	0.082
Students work on KA about 1x per week in School	156	0.17	0.18	0.013
Students work on KA more than 1x per week in School	156	0.16	0.11	-0.048
Student KA use after School				
Students Never work on KA after School	156	0.63	0.65	0.028
Students work on KA less than 1x per week after School	156	0.28	0.29	0.002
Students work on KA about 1x per week after School	156	0.07	0.04	-0.032
Students work on KA more than 1x per week after School	156	0.02	0.02	0.002
Student non-KA CAT use in School				
Students never work on non-KA CAT in school	156	0.13	0.10	-0.029
Students work on non-KA CAT less than 1x per week in School	156	0.11	0.11	0.002
Students work on non-KA CAT about 1x per week in School	156	0.14	0.12	-0.014
Students work on non-KA CAT more than 1x per week in School	156	0.63	0.67	0.041
Student non-KA CAT use after School				
Students Never work on non-KA CAT after School	156	0.26	0.24	-0.024
Students work on non-KA CAT less than 1x per week after School	156	0.26	0.28	0.020
Students work on non-KA CAT about 1x per week after School	156	0.21	0.21	0.002
Students work on non-KA CAT more than 1x per week after School	156	0.27	0.27	0.002

Notes: Percentages represent student averages among Grade-School units (the unit of randomization).

BACK

AISD: Main Results by Grade

	No Controls	Grade FEs	Grade FEs/Student Controls
All Grades	0.036	0.044	0.025
	(0.092)	(0.083)	(0.076)
N	10,979	10,979	10,979
Grade 3	0.223*	0.223*	0.106
	(0.128)	(0.128)	(0.090)
N	1,600	1,600	1,600
Grade 4	0.102	0.102	0.096
	(0.142)	(0.142)	(0.110)
N	1,638	1,638	1,638
Grade 5	0.133	0.133	0.104
	(0.137)	(0.137)	(0.115)
N	1,694	1,694	1,694
Grade 6	0.218	0.218	0.161
	(0.145)	(0.145)	(0.124)
N	2,302	2,302	2,302
Grade 7	-0.103	-0.103	-0.045
	(0.151)	(0.151)	(0.113)
N	1,785	1,785	1,785
Grade 8	-0.293	-0.293	-0.324
	(0.362)	(0.362)	(0.359)
N	1,960	1,960	1,960

Notes: OLS regressions of standardized 2022 Math STAAR scores on treatment. Standard errors clusted at the grade/school level. Controls for III include: age, sex, race, ethnicity, days missed, english learner status, special ed status, free lunch eligibility.

BACK

AISD: Subgroup Analysis Contd. (Grades 3-6)

		Special Ed			Economic Indicator		
Specification	All	Spec-Ed	Not	F-Stat	Free Lunch	Not	F-Stat
No Controls	0.171**	0.169**	0.187**	0.06	0.128**	0.189**	0.88
	(0.069)	(0.078)	(0.075)		(0.062)	(0.084)	
	7234	1044	6190		3521	3713	
Grade FEs	0.172**	0.160**	0.190**	0.05	0.127**	0.191**	0.82
	(0.070)	(0.080)	(0.076)		(0.063)	(0.085)	
	7234	1044	6190		3521	3713	
Grade FEs w/ Controls	0.122**	0.116	0.122**	0.00	0.110*	0.133*	0.81
	(0.058)	(0.071)	(0.061)		(0.057)	(0.071)	
	7234	1044	6190		3521	3713	

Notes: OLS regressions of standardized 2022 Math STAAR scores on treatment, broken out by subgroup. Standard errors clusted at the grade/school level. Controls for III include: age, sex, race, ethnicity, days missed, english learner status, special ed status, free lunch eligibility. F-Stat comparing treatment coefficient from both subgroups.

BACK

AISD: OLS relationship, class averages

	I	II	III
25-35 minutes	-0.240***	-0.389***	-0.200
	(0.090)	(0.093)	(0.195)
N	7,915	7,915	7,915
35+ minutes	-0.010	0.048	0.190
	(0.082)	(0.132)	(0.172)
N	7,915	7,915	7,915
2-3 Level-ups	0.142	0.330**	0.168
	(0.149)	(0.163)	(0.260)
N	7,915	7,915	7,915
3+ Level-ups	0.243***	0.065	0.172
	(0.060)	(0.110)	(0.128)
N	7,915	7,915	7,915

Notes: OLS regression of class average practice (leave out individual) on standardized 2022 Math STAAR scores. Does not include grade 3, as students do not test in grade 2. I: Controls for student demographics and 2021 STAAR Math Scores, Grade FEs. SEs clustered at the grade-school level. II: Adds School FEs. III: Adds Grade-School FEs.

BACK

AISD: Onboarding Survey and Practice (grade 7-8)

	Coeff	SE
Teacher attended PD session during district PD Day	8.29**	(3.03)
Teacher held first meeting with khollaborator	2.88	(2.90)
Teacher held followup meeting with khollaborator prior to Oct 1st, 2021	3.44	(3.03)
How many days per week will your students receive math instruction?	3.16	(2.91)
About how long is each math class?	0.40**	(0.15)
How easy do you expect it would be for you to provide some time each week to work on Khan Academy DURING CLASS?		
I think I can provide some practice time once a week, for less than 30 minutes	-4.80*	(2.52)
I think I can provide some practice time once a week, for more than 30 minutes	6.19	(4.27)
I think I can provide some practice time more than once a week, for less than 30	-6.85**	(2.94)
I think I can provide some practice time more than once a week, for more than 30	5.34	(5.37)
How easy do you expect it would be for students to practice on Khan Academy AFTER SCHOOL for at least an hour a week?		
Easy	1.53	(3.33)
I don't usually assign my students any homework	0.63	(3.14)
I only assign my students homework if they are unable to complete practice assig	6.48	(4.50)
I usually assign homework each week	1.99	(3.16)
Most students don't often complete their homework	2.61	(3.48)
I make homework a graded requirement (even just a participation grade)	4.04	(3.97)
Some students do not have access to use or borrow a computer or tablet for homew	-6.04*	(2.99)
Students have access to a computer lab at school that they can use during lunch	-1.09	(3.78)
Which homework guideline do you think would work for your students (and their parents) to try to follow?		
What practice time you don't finish during school, finish at home.	-3.30	(3.55)
Once a week, focus on practicing for at least an hour	1.72	(3.28)
Two times a week for at least 30 minutes each time	0.95	(3.49)
Try a little every day	5.01	(3.06)
Do at least an hour of practice, but see if you can do even more.	0.00	(0.00)
How easy do you expect it would be to communicate with parents about the program, asking for help to facilitate at least an hour of practice at home?		
Most parents read messages when I send them by email	-2.02	(3.33)
Most parents read messages when I send them by text or use Remind	-4.90	(3.45)
Most parents read messages when I provide a handout for students to take home to	-0.34	(3.98)
I have used the app Remind in the past to message parents	-2.44	(2.26)
Most parents are too busy or too overwhelmed to help	3.44	(3.49)
Most parents do not want their children assigned homework	-3.42	(2.80)

Notes: OLS coefficients of a teacher's average weekly practice time on their answers to onboarding questions. SE clustered at grade student level, FEs for grade. N: 28 teachers (treatment, grade 7-8)

BACK

AISD: Prior Experience and Practice (grades 3-6)

	Coeff	SE
Percent of Students with Access to Tech	-0.04	(0.16)
Teaches multiple grades	-26.91***	(5.28)
Prior Khan Use		
Never Used KA	2.31	(9.41)
Use KA for Demonstration Only	-5.25	(7.56)
Use KA for Optional Practice	4.94	(10.83)
Use KA for Required Practice	0.82	(10.60)
Student Independent Work in School		
Students not Expected to Practice Math Indep. in School	0.00	(0.00)
Students Expected to Practice Math Indep. in School 1-30 mins weekly	-10.06	(6.18)
Students Expected to Practice Math Indep. in School 31-60 mins weekly	-8.41	(8.44)
Students Expected to Practice Math Indep. in School 61-90 mins weekly	15.16	(9.47)
Students Expected to Practice Math Indep. in School 90 plus mins weekly	2.99	(7.97)
Student Independent Work after School		
Students not Expected to Practice Math After School	-8.86	(12.98)
Students Expected to Practice Math After School 1-30mins weekly	1.44	(7.65)
Students Expected to Practice Math After School 31-60mins weekly	1.59	(7.89)
Students Expected to Practice Math After School 61-90mins weekly	7.35	(10.12)
Students Expected to Practice Math After School 90 plus mins weekly	-14.20	(14.79)
Khan Academy Use in School		
Students Never work on KA in school	-5.62	(8.84)
Students work on KA less than 1x per week in School	11.68	(8.12)
Students work on KA about 1x per week in School	-8.00	(8.01)
Students work on KA more than 1x per week in School	-2.39	(11.01)
Khan Academy Use after School		
Students Never work on KA after School	1.08	(7.56)
Students work on KA less than 1x per week after School	4.91	(7.45)
Students work on KA about 1x per week after School	-12.93*	(6.98)
Students work on KA more than 1x per week after School	-27.23***	(7.94)
Non-Khan CAL Use in School		
Students never work on non-KA CAT in school	3.44	(11.37)
Students work on non-KA CAT less than 1x per week in School	16.32	(21.39)
Students work on non-KA CAT about 1x per week in School	12.81	(10.32)
Students work on non-KA CAT more than 1x per week in School	-14.16*	(8.27)
Non-Khan CAL Use after School		
Students Never work on non-KA CAT after School	-2.31	(8.49)
Students work on non-KA CAT less than 1x per week after School	-0.25	(9.21)
Students work on non-KA CAT about 1x per week after School	5.21	(8.64)
Students work on non-KA CAT more than 1x per week after School	-1.57	(5.91)

Notes: OLS coefficients of a teacher's average weekly practice time on their answers to a pre-randomization survey. SE clustered at grade student level, FEs for grade. N: 84 teachers (treatment, grades 3-8)

BACK

AISD: Prior Experience and Practice (grades 7-8)

	Coeff	SE
Percent of Students with Access to Tech	0.08	(0.09)
Teaches multiple grades	10.81**	(3.65)
Prior Khan Use		
Never Used KA	-4.30	(3.44)
Use KA for Demonstration Only	1.88	(4.44)
Use KA for Optional Practice	-1.57	(2.49)
Use KA for Required Practice	8.68	(5.45)
Student Independent Work in School		
Students not Expected to Practice Math Indep. in School	0.00	(0.00)
Students Expected to Practice Math Indep. in School 1-30 mins weekly	0.03	(2.80)
Students Expected to Practice Math Indep. in School 31-60 mins weekly	-1.44	(2.81)
Students Expected to Practice Math Indep. in School 61-90 mins weekly	4.17	(3.27)
Students Expected to Practice Math Indep. in School 90 plus mins weekly	-5.28	(3.63)
Student Independent Work after School		
Students not Expected to Practice Math After School	-0.78	(3.48)
Students Expected to Practice Math After School 1-30mins weekly	0.24	(2.59)
Students Expected to Practice Math After School 31-60mins weekly	-2.37	(3.16)
Students Expected to Practice Math After School 61-90mins weekly	-1.81	(4.24)
Students Expected to Practice Math After School 90 plus mins weekly	24.20***	(0.23)
Khan Academy Use in School		
Students Never work on KA in school	0.59	(3.87)
Students work on KA less than 1x per week in School	0.35	(3.56)
Students work on KA about 1x per week in School	-3.75	(2.78)
Students work on KA more than 1x per week in School	0.00	(0.00)
Khan Academy Use after School		
Students Never work on KA after School	-0.57	(4.14)
Students work on KA less than 1x per week after School	-3.46	(2.52)
Students work on KA about 1x per week after School	4.34	(5.66)
Students work on KA more than 1x per week after School	0.00	(0.00)
Non-Khan CAL Use in School		
Students never work on non-KA CAT in school	-3.33	(2.24)
Students work on non-KA CAT less than 1x per week in School	-3.49	(3.44)
Students work on non-KA CAT about 1x per week in School	-2.43	(2.19)
Students work on non-KA CAT more than 1x per week in School	5.63	(3.26)
Non-Khan CAL Use after School		
Students Never work on non-KA CAT after School	-2.97	(2.50)
Students work on non-KA CAT less than 1x per week after School	1.20	(3.39)
Students work on non-KA CAT about 1x per week after School	1.64	(7.56)
Students work on non-KA CAT more than 1x per week after School	1.06	(4.14)

Notes: OLS coefficients of a teacher's average weekly practice time on their answers to a pre-randomization survey. SE clustered at grade student level, FEs for grade. N: 28 teachers (treatment, grades 7-8)

BACK