MATH TALK, PUT TO THE TEST



See OKO in action: okolabs.ai/video

> "Do you know how huge it is to have all your students talking about math, doing math, working in groups? Every kid was engaged."
>
> Hamilton County teacher, interviewed in the independent study below

THE STUDY

OKO is AI-powered small-group math for grades 3–12. WestEd ran a randomized controlled trial of it in 11 Hamilton County, Tennessee schools: 20 fourth- and fifth-grade teachers and 1,040 students, randomized by classroom to OKO or their usual software.

Jan–Apr 2026, avg. 11.2 weeks. Attrition 4.04%: low under What Works Clearinghouse standards, per WestEd.

INDEPENDENT EVALUATOR

WestEd

FUNDED BY

IES Institute of Education Sciences · Gates Foundation · Accelerate

WHAT IT FOUND

RANDOMIZED COMPARISON · ALL STUDENTS

OKO classrooms showed more math growth than comparison classrooms, though with only 20 teachers the difference wasn't statistically conclusive.

Growth above projection +1.46 RIT* (p = 0.166); spring score +1.40 (p = 0.218); threshold p < 0.10.

EXPLORATORY

~1.5 MONTHS

EXTRA MATH GROWTH*

Students who used OKO gained approximately 1.5 months of extra math growth in ten weeks.

Teachers chose who used OKO, so this isn't causal. +2.46 RIT on growth (p = 0.033) and spring score (p = 0.047).

CLASSROOM OBSERVATIONS

MORE MATH TALK

STUDENT TO STUDENT

OKO students explained their thinking and corrected each other's errors. Comparison students were on task but rarely talked math with each other; their questions went to the teacher.

Students were engaged in 9 of 10 OKO classrooms observed.

TEACHER-REPORTED

2.8 HOURS

SAVED EACH WEEK

OKO teachers estimated it saved them 2.8 hours a week; comparison teachers, 1.0 from their usual software.

Most saved on: facilitating math discussion. Post-survey, 10 vs. 8 teachers.

OKO'S IN-PLATFORM MEASURE

Collaboration scores also climbed steadily across the 255 students using OKO (0.03 points a week, p < 0.001); validation underway.

* RIT: the score scale of NWEA's MAP Growth test. Growth above projection: spring score minus NWEA's projection from the winter score. ~1.5 months: the +2.46 RIT gap over control students, set against NWEA's typical winter-to-spring growth (an estimate).

SMALL GROUPS, MEGA MINDS

WHAT IT LOOKED LIKE

CLASSROOM OBSERVATIONS

WestEd observed 14 classrooms and interviewed teachers in both conditions.

WHAT OBSERVERS SAW

Peer help. Students talked each other through problems. A teacher heard: "I don't know how to do that. Could you show me how to do this?"

Persistence. Students tried multiple strategies and worked to help stuck teammates. In two classrooms, students wanted to keep going after OKO time ended.

WHAT OKO TEACHERS SAID

Quieter voices. One teacher said students who usually don't talk in class sometimes did in OKO. "I was able to give my quieter, shier students shout-outs."

Enthusiasm. Four of six interviewed teachers said students were always excited to use OKO. "My kids never complained. Not one time."

IN COMPARISON CLASSROOMS

Engagement ran through the teacher (call-and-response, popsicle sticks to pick who answers), and students rarely talked math with each other.

On their usual software, students were "going through the motions to complete it." Comparison teacher

ESSA EVIDENCE

LXD Research
Tier III: 2026 WestEd trial
Tier IV: OKO logic model

Tier III rests on WestEd's adjusted comparison of students who used OKO with control students: +2.46 RIT on both growth above projection ($p = 0.033$) and spring score ($p = 0.047$). Teachers chose those students, who started higher. On the same two measures, the randomized comparison ($p = 0.166$, 0.218) and a matched comparison ($p = 0.118$, 0.129) were not significant.

LEANLAB EDUCATION
Tiers III and IV:
earlier WestEd study

A separate WestEd study, peer-reviewed (CSCL 2025): 299 students over 2–3 weeks, no comparison group. Fractions scores ($p = 0.006$) and math anxiety ($p = 0.019$) improved. Both studies: okolabs.ai/research

RESEARCH PARTNERS

INDEPENDENT EVALUATOR

WestEd

Designed and ran the trial and every analysis.

INDEPENDENT EVIDENCE REVIEWERS

LXD RESEARCH at Charles River Media · LEANLAB EDUCATION

Rated OKO's evidence against ESSA tiers.

RESEARCH COLLABORATORS

EdTechnical · measurement science
Digital Promise · Gates-funded ENCORE initiative
Joan Ganz Cooney Center · co-design

THE TRIAL PAPER

Feng, Lee, Hsu, Miller & Marston (2026). EdWorkingPaper 26-1656, under peer review.
edworkingpapers.com/ai26-1656

WHAT COMES NEXT

We keep investing in R&D. Independent validation of our collaboration measure is underway, and a larger, fully powered trial is in design.

PILOT OKO IN YOUR SCHOOL

Let's scope a pilot with your teachers. We're also happy to share the full WestEd report and our logic model. Or watch OKO in action at okolabs.ai/video.