The design decides the result.

Harvard researchers found AI tutoring more than doubled learning gains. Penn researchers found an unrestricted chatbot left students worse off. The difference was the design — and design is what ArthurAI™ is built on.

UNRESTRICTED ANSWERS
–17% on the exam once AI was removed
Penn researchers, high school maths

→

GUARDRAILS ADDED
Harm largely avoided — no exam gain
Same study, hint-only tutor

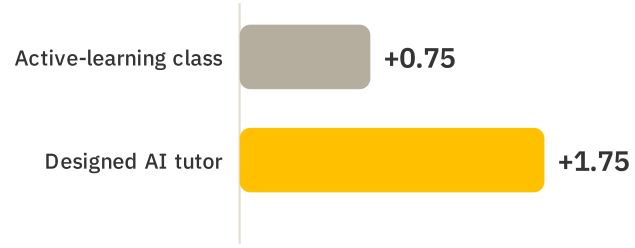
→

BUILT TO TEACH
More than double the median learning gain
Harvard researchers, university physics

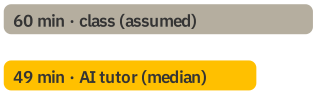
Independent research — not an ArthurAI trial. *Each result is from the study named; none used ArthurAI.*

HARVARD RESEARCHERS · SCIENTIFIC REPORTS, 2025

A well-designed AI tutor beat an active-learning class



Median gain over pre-test median (2.75), test points



Time on task — AI tutor: measured median; class: 60 min assumed by the authors

4.1
engagement
vs 3.6 in class

3.4
motivation
vs 3.1 in class

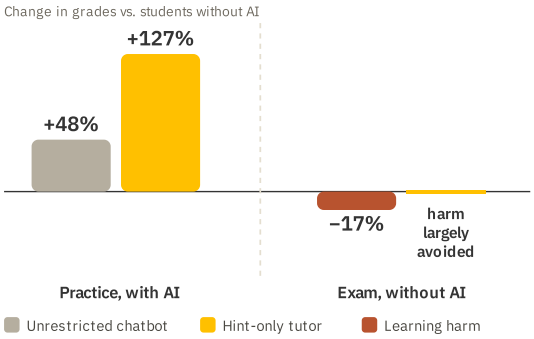
83%
rated the tutor's explanations as good as or better than instructors'

Crossover RCT: 194 students in an introductory physics course, each taught one lesson each way. Effect 0.73–1.3 SD (quantile regression). Class time on task assumed at 60 min by the authors.

Independent research — not an ArthurAI trial.

PENN RESEARCHERS · PNAS, 2025

Answers on demand helped practice — then hurt the exam



Field experiment with nearly a thousand high-school maths students. When AI was taken away, the unrestricted group scored 17% lower than students who never had it; a tutor built to give hints, not answers, largely mitigated that harm but did not raise exam scores.

Independent research — not an ArthurAI trial.

"Having access to AI is not the same as being educated by AI."

The Harvard authors are explicit: an AI tutor *"should not replace in-person teaching"* — it brings every student to the point where class time does the most good. That is the ArthurAI motto in practice: **the AI reasons; the educator decides.**

Figures verified against original papers, November 2026.

ArthurAI™ and MindHYVE.ai™ are trademarks of MindHYVE.ai, Inc. ArthurAI is operated by Eve-Education, LLC, a wholly-owned subsidiary of MindHYVE.ai, Inc.
© 2026 Eve-Education, LLC. arthurgrid.ai · hello@mindhyve.ai · +1 (949) 200-8668 · Newport Beach · Islamabad · Nairobi

Beyond one study: the pattern holds where design is deliberate

WORLD BANK RESEARCHERS · NIGERIA

0.23 SD in English

Six weeks of after-school AI tutoring for first-year senior secondary students. 0.31 SD on the composite test.

Caveat: heavy, unequal attrition; excluded from the meta-analysis.

STANFORD RESEARCHERS

+4 pp
mastery, all students

+9 pp
with lower-rated tutors

AI that coaches the *human* tutor in real time lifted topic mastery — most where tutors were least experienced. (Preprint.)

WORLD BANK RESEARCHERS · 2026

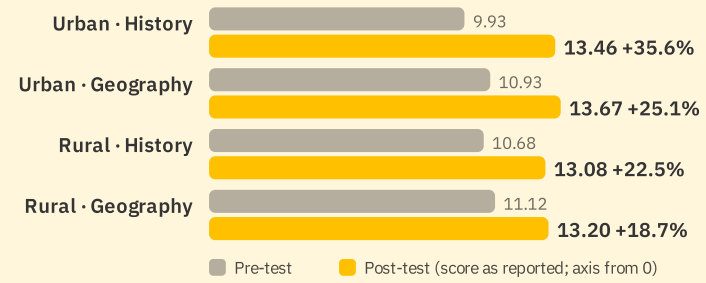
0.125 SD average

Meta-analysis of 14 randomised trials in ten economies: adaptive and AI tools helped, modestly on average. Newer generative tools did not beat earlier adaptive software.

The average hides the spread: design decides.

Independent research — not an ArthurAI trial. *Effect sizes come from different tests, ages and durations and are not directly comparable.*

OUR OWN PILOT · ARTHURAI SLE · ISLAMABAD, 2026



In our government-supervised pilot in Islamabad, over 80% of students improved on independently designed tests. We haven't yet run a controlled trial — your pilot can include one.

Class 8 History and Geography in two public schools (one urban, one rural), under Federal Directorate of Education oversight. Pre- and post-tests independently designed by the Pakistan Institute of Education.

Limits: Pre/post only, no comparison group, about 103 students, four weeks. It shows improvement while using ArthurAI — not that ArthurAI caused it.

mindhyve.ai/research/case-studies/fde-pakistan-pilot/

Every ArthurAI pilot includes a learning measure at no extra charge, co-designed with your faculty: pre/post learning gain · time on task · engagement & motivation · 30-day retention. Aggregate results, published only with the methodology and your written consent.

THE HARVARD TUTOR'S PRACTICES, AND WHERE ARTHURAI PUBLISHES THE SAME

Six of seven are published on arthurgrid.ai; growth mindset is not claimed.

- **Accuracy** — every claim cites its lesson or textbook page (PLE · SLE)
- **Scaffolding** — sequenced steps that cannot be skipped (PLE)
- **Cognitive load** — a cognitive-load-aware design commitment (Why 06)
- **Active learning** — Socratic questions, not answers, in group work (SLE)
- **Targeted practice** — drills on each learner's weak skills (PLE)
- **Self-pacing** — no term dates; resumes where the learner stopped (PLE)

WHAT THESE STUDIES DO NOT SHOW

- The independent studies did not test ArthurAI; our own pilot is shown separately above.
- Harvard: two first-exposure lessons in one course.
- Penn: the hint tutor avoided harm but did not raise exam scores.
- Nigeria: heavy, unequal attrition.
- Stanford: a preprint; AI assisted human tutors.
- World Bank meta: a modest average, from a small number of studies.

SOURCES — READ THE ORIGINAL PAPERS

Kestin et al. AI tutoring outperforms in-class active learning. Sci Rep 15:17458 (2025). doi.org/10.1038/s41598-025-97652-6
Bastani et al. Generative AI without guardrails can harm learning. PNAS 122(26) (2025). doi.org/10.1073/pnas.2422633122
De Simone et al. From Chalkboards to Chatbots. World Bank PRWP 11125 (2025). openknowledge.worldbank.org/entities/publication/15e1ff08-15ae-4f7a-b2a8-d146e6c113ee
Wang et al. Tutor CoPilot. arXiv 2410.03017 (preprint). arxiv.org/abs/2410.03017
Burneo et al. Can EdTech Close Learning Gaps? WDR 2026 background paper (2026). thedocs.worldbank.org/.../Can-EdTech-Close-Learning-Gaps.pdf
Our pilot: ArthurAI Pakistan Pilot Overview, V1.0 (MindHYVE.ai, 2026). mindhyve.ai/research/case-studies/fde-pakistan-pilot/

Charts drawn by ArthurAI from reported figures; none reproduced. Affiliations imply no endorsement.

Figures verified against original papers, November 2026.

ArthurAI™ and MindHYVE.ai™ are trademarks of MindHYVE.ai, Inc. ArthurAI is operated by Eve-Education, LLC, a wholly-owned subsidiary of MindHYVE.ai, Inc.
© 2026 Eve-Education, LLC. arthurgrid.ai · hello@mindhyve.ai · +1 (949) 200-8668 · Newport Beach · Islamabad · Nairobi