Putting AI to the Test — An 8,647-Pupil RCT in Rural Sichuan Finds AI-Personalised Maths Practice No Better Than Non-Personalised Software

AI in Education · Good practice · China

Quality score: 47/100

Evidence strength: Moderate

Summary

Large three-arm RCT in 315 rural Sichuan primary classes: AI-personalised maths practice did no better than identical non-personalised software or usual activities (about –0.04 SD), with estimates precise enough to rule out non-trivial gains.

Key dimensions

Dimension	Score
Evidence of learning impact	1/3
Equity & inclusion	2/3
Transparency, ethics & data protection	2/3
Teacher capability & pedagogy	1/3
Scalability & sustainability	1/3

Evaluated by C-NAPSE

Data sources

C-NAPSE web research — accessed 2026-11-12

Stanford Rural Education Action Program (REAP) research team (Ma, Feng, Fairlie, Liu, Loyalka, Rozelle, Zhang) — accessed 2026-11-12

Stanford FSI — Putting AI to the Test (Economics of Education Review) — accessed 2026-11-12

CESifo Working Paper 12837 (EconPapers) — accessed 2026-11-12

IZA DP 19025 — Making AI tutoring work for poorer schools (Dang) — accessed 2026-11-12

Cite this — Evidoria. *Putting AI to the Test — An 8,647-Pupil RCT in Rural Sichuan Finds AI-Personalised Maths Practice No Better Than Non-Personalised Software*. Persistent ID: c6e89b4d-d8a2-4d5f-b733-0e2a0765a6bf.