How the ecosystem is supposed to help

Three proposed pathways. Each link needs local evidence; none implies inevitable social change.

	WHAT WE WOULD OFFER	WHY IT MIGHT HELP	WHAT TO OBSERVE	LONGER-TERM AIM
LEARN	**Kosplora + companion** Practice, feedback and relevant sources.	**Usable understanding** Attempt, explanation and correction.	**Independent skills** A new task, with chosen accessibility supports.	**More capable people** Not just more completed lessons.
COOPERATE	**Community + civic** Find collaborators; clarify decisions.	**Feasible shared work** Consent, resources and named responsibilities.	**Useful work delivered** Follow-through, dissent and who bears the work.	**Collective agency** Not just more meetings.
STEWARD	**Needs + resources** Compare repair, sharing and services.	**Better-fitted options** Life-cycle costs and maintenance made visible.	**Services that last** Reliability, total burden and maintainers.	**Dignity + ecology** Not just more purchases.

CONDITIONS: time • access • resources • legitimate authority • privacy • maintenance

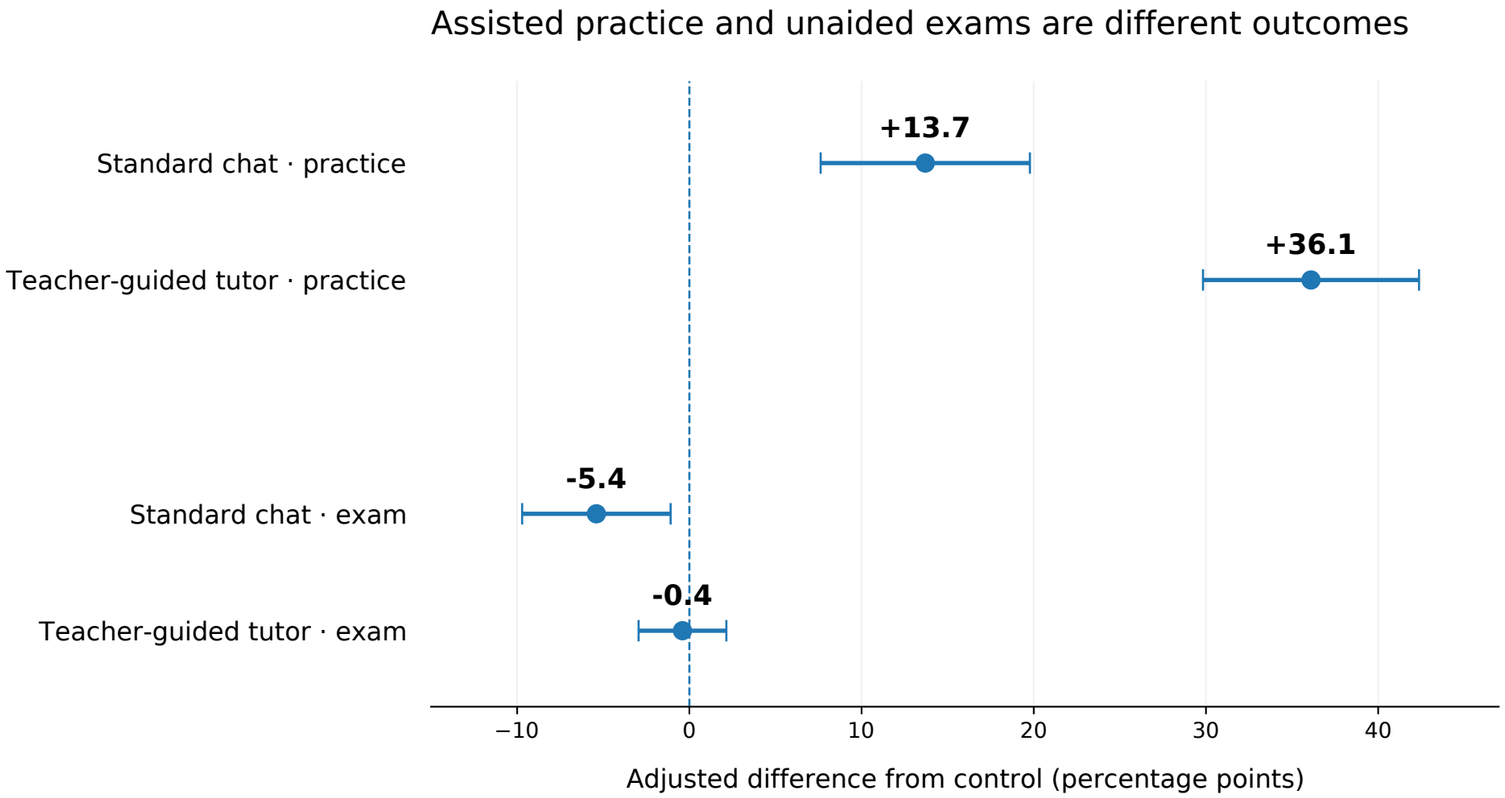
FAILURE SIGNAL

More activity without better options, capabilities or conditions is not the intended outcome.

Basis: the Futurism Institute’s proposed projects and plan. An original synthesis, not a measured impact model.
Dashed arrows = hypotheses. Wider outcomes also depend on policy, income, institutions and local conditions.

More help is not the same as more learning

A field experiment separated performance with AI from a subsequent exam without it.



STUDY CONTEXT

One Turkish high school; grades 9–11.

Main sample: 839 students. Regression: 2,848 student-session observations per outcome.

READ THE SCALE

Percentage points, not relative percent gains. Zero means no estimated difference from control.

IMPLICATION FOR KOSPLORA

Measure transfer separately from assisted completion. The tutor exam result does not establish a benefit.

[S1] Bastani et al. (2025), PNAS, Table 1. Effects = published coefficients × 100.
Whiskers: approximate 95% normal intervals, estimate ± 1.96 × reported class-clustered SE; not author-reported intervals.

The design matters more than the “AI tutor” label

Two studies answer different questions. Their findings should not be pooled into one improvement rate.

[S1] HIGH-SCHOOL MATHEMATICS

Assistance during practice

Standard chat helped assisted work but reduced the subsequent unaided exam score. Guarded tutoring did not show an exam benefit.

Class-randomized study. One school; short-term outcomes.

[S2] UNDERGRADUATE PHYSICS

A structured instructional package

A designed AI tutor produced higher immediate post-test scores than active classroom instruction for two lessons.

Harvard crossover study; 194 eligible students. AI instruction also differed in pacing and setting.

NOT ESTABLISHED BY THESE RESULTS

Universal superiority / Durable real-world transfer / Institute-specific impact

DESIGN IMPLICATION

Test the complete learning arrangement against a credible alternative, rather than against nothing.

[S1] Bastani et al. (2025), PNAS. [S2] Kestin et al. (2025), Scientific Reports, Results and Methods.

Selected studies, not a systematic review. Different learners, tasks, designs and endpoints prevent a pooled effect claim.

Test the world and the companion separately

A proposed 2 × 2 study asks whether either adds value beyond the same lesson in a simpler interface.

NO AI COMPANION	AI COMPANION
A **Conventional lesson** Same topic, resources and assessment.	**B** **Lesson + companion** Add structured AI guidance.
C **Simulated world** Explore the same concepts through a model.	**D** **World + companion** Combine simulation and structured guidance.

ALL FOUR TAKE THE SAME TESTS

New task outside the game

Delayed retention; blinded scoring where feasible; chosen accessibility supports retained.

Also record enjoyment, withdrawal, access barriers, staff effort and total cost.

Compare B–A, C–A and the interaction (D–C)–(B–A).

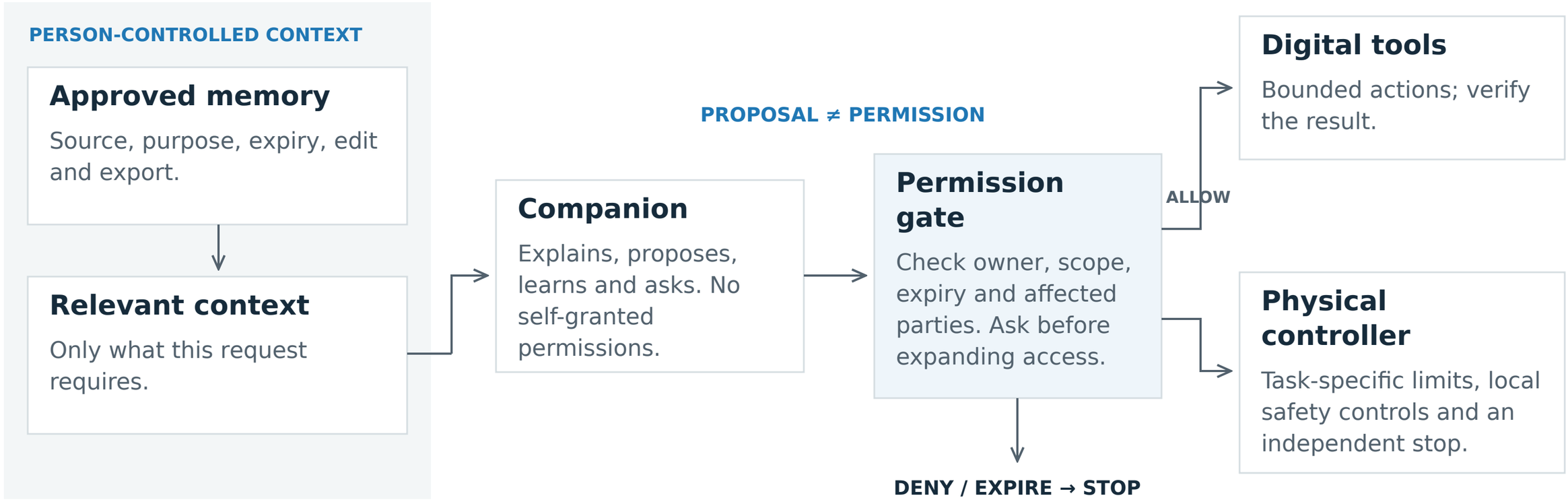
KEEP THE INTERPRETATION NARROW

Minecraft agent performance is not evidence that a simulated world teaches humans better.

Design proposal. [S1–S2] motivate transfer testing; [S6] is agent research, not a learner trial.
Predefine the protocol, comparison and meaningful effect. Use a feasibility pilot first; set sample size through a power analysis.

A companion may suggest more than it may do

The person, the language model, the permission service and the physical controller have different authority.



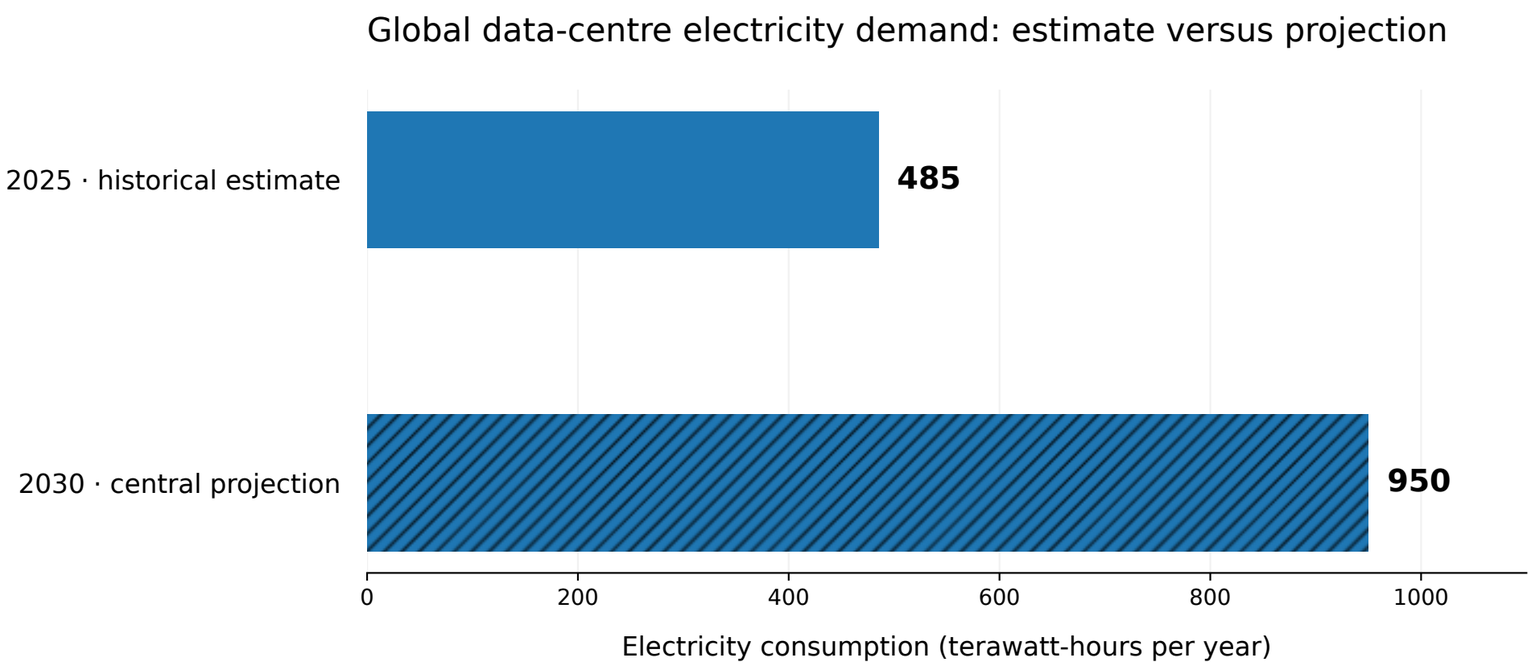
ALWAYS AVAILABLE

Inspect actions • revise permissions • revoke access • export memory • use a non-AI route

Original architecture proposal. [S4] NIST SP 800-207A informs authorization boundaries; [S5] motivates layered physical safeguards. Boxes describe required responsibilities. They are not implemented components, a threat model or a safety certification.

Compute savings have to be measured

Global sector demand gives context. It cannot establish the footprint or savings of a particular companion.



SCOPE

All data centres worldwide, not AI alone.

FOR AI SEARCHES

Compare end-to-end work: generation, retrieval, checking, storage and correction.

No savings result is claimed here.

USE THE RIGHT DENOMINATOR

Measure resources per useful, checked outcome. Report emissions only with an explicit accounting method.

[S3] IEA, Key Questions on Energy and AI, 16 April 2026, Executive summary. Original chart from reported values; CC BY 4.0.
2025 = historical estimate; 2030 = central projection. No confidence interval, annual path, AI-only share or per-query estimate inferred.

Where the next evidence is needed

A research ledger for the ecosystem. “Not supplied” is a boundary of this packet, not a claim that no research exists.

AREA	CLAIM TO TEST	EVIDENCE TO COLLECT	IF IT FAILS
Kosplora + learning world	Knowledge transfers beyond the lesson.	Novel and delayed tasks; access barriers.	Redesign the instruction.
Private companion	Continuity helps without unwanted dependence.	Longitudinal outcomes; correction and exit tests.	Narrow memory or initiative.
AI Searches	Reuse saves resources without carrying errors forward.	Matched tasks; applicability errors; full cost.	Restrict reuse conditions.
Civic + community tools	Assistance improves decisions while preserving dissent.	Participant influence; consent; actual follow-through.	Simplify or remove the tool.
Physical interfaces	A body adds useful, safely bounded assistance.	Task benefit; hazards; stop and recovery tests.	Remain digital or stationary.
Places + material systems	Shared services remain useful and maintainable.	Reliability; burdens; paid maintenance; handoff.	Change scope or ownership.

EVIDENCE BOUNDARY

No Institute impact dataset was supplied. Test one bounded pathway before attributing ecosystem-wide benefits.

Original prioritization from the Institute’s plan. The wider aims (culture, dignity, agency and ecology) remain explicit values.
This ledger is selective, not an exhaustive literature review or an assessment of which projects are operational.

Sources, calculations and reading rules

The numbers can be traced. The proposed relationships can be challenged. Neither is presented as a finished proof.

S1 / Bastani et al. / 2025

Generative AI without guardrails can harm learning

Table 1; author manuscript pp. 10 and 33

S2 / Kestin et al. / 2025

AI tutoring outperforms in-class active learning

Results: Learning gains; Methods: Study design

S3 / International Energy Agency / 2026

Key Questions on Energy and AI

Executive summary; published 16 April 2026

S4 / NIST / 2023

A Zero Trust Architecture Model for Access Control

SP 800-207A

S5 / Google DeepMind / Accessed 2026

Responsibly advancing AI and robotics

Physical safety; responding to safety distances

S6 / Wang et al. / 2023

Voyager: An Open-Ended Embodied Agent with LLMs

Components and experiments

CALCULATION & SCOPE

Learning: 100 × coefficient; interval = 100 × (coefficient ± 1.96 × clustered SE).
No pooled study effect. No invented Institute outcomes. No research evidence attributed to concept diagrams.

Exact URLs, source locators, extracted values and longer limitations are included in the HTML page and editable source package.
Selected primary sources checked for this atlas; not a systematic review. No website deployment, safety audit or accessibility certification claimed.

Connect the projects through a real task

Illustrative pilot: a group wants an accessible shared room for learning and making.

	WHAT PEOPLE DO	SUPPORTING SYSTEM	RECORD TO INSPECT
1	**Define the need** Who needs the room, and what is currently missing?	**Needs + community**	Need, constraints, alternatives
2	**Learn and explore** Try layouts; inspect what the simulation assumes.	**Kosplora + learning world**	Rule version + open questions
3	**Compare options** Include repair, sharing and using an existing venue.	**Conscious Consuming**	Options + full maintenance cost
4	**Agree on a bounded test** People consent to tasks; the venue approves use.	**Civic tools + companion**	Authority, scope + dissent
5	**Try it in reality** Test a reversible layout; do not rely on the game score.	**Local participants**	Observed use + access barriers
6	**Maintain, revise or stop** Name a maintainer and publish only the agreed record.	**Stewardship + AI Searches**	Handoff + corrections

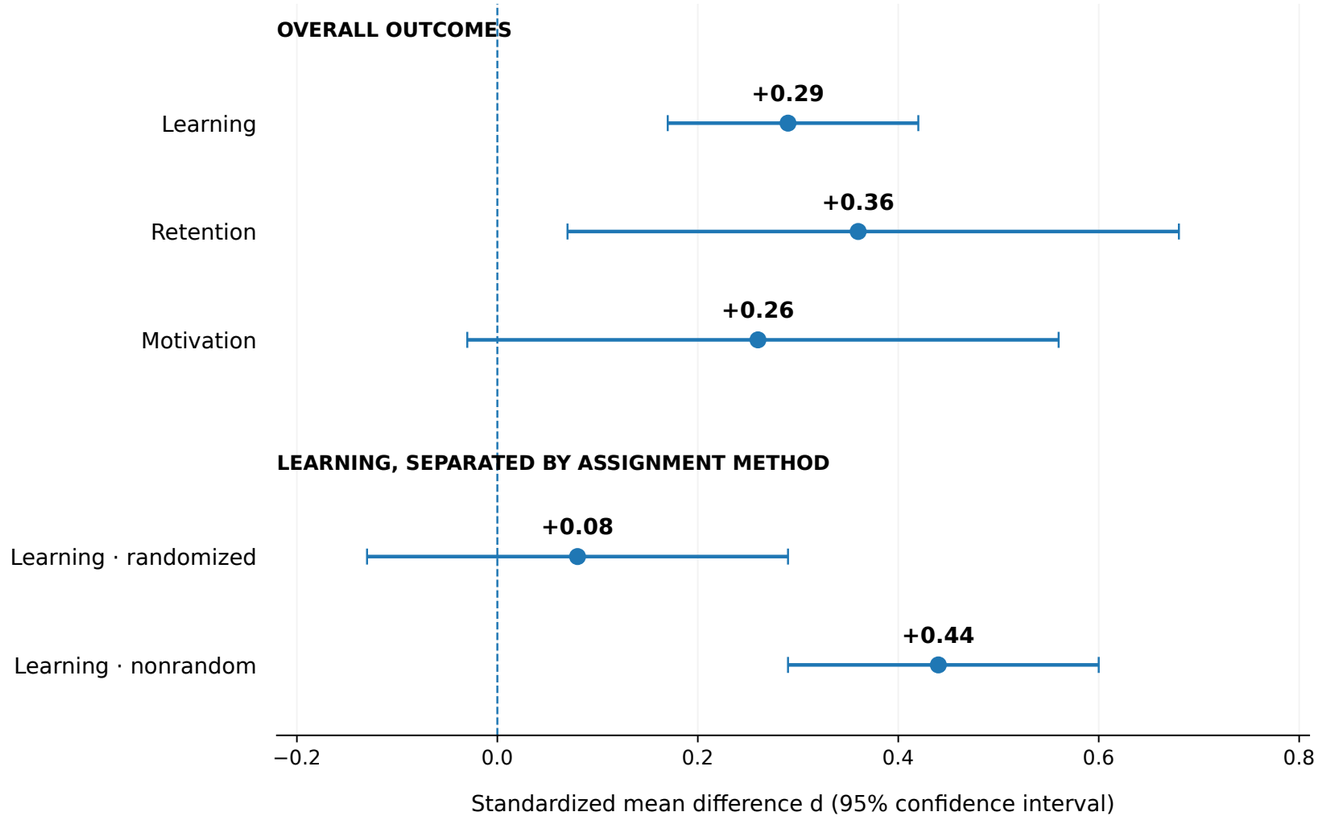
NOT EVERY ROUTE ENDS IN A NEW INSTITUTION

A valid outcome is to use something that already works, or to decide not to build.

Original design synthesis. Dashed links show a proposed sequence, not measured causal effects.
No venue, partnership, participant group, budget or completed pilot is asserted.

Whether games help is worth testing

A historical synthesis found positive averages; its randomized learning subset was less conclusive.



READ THIS AS

Reasons to run a careful local test.

1990–2012 research. Mixed games, learners and methods. Not an AI-world or Minecraft trial.

ZERO IS IMPORTANT

Intervals crossing zero do not establish an advantage, or prove no effect.

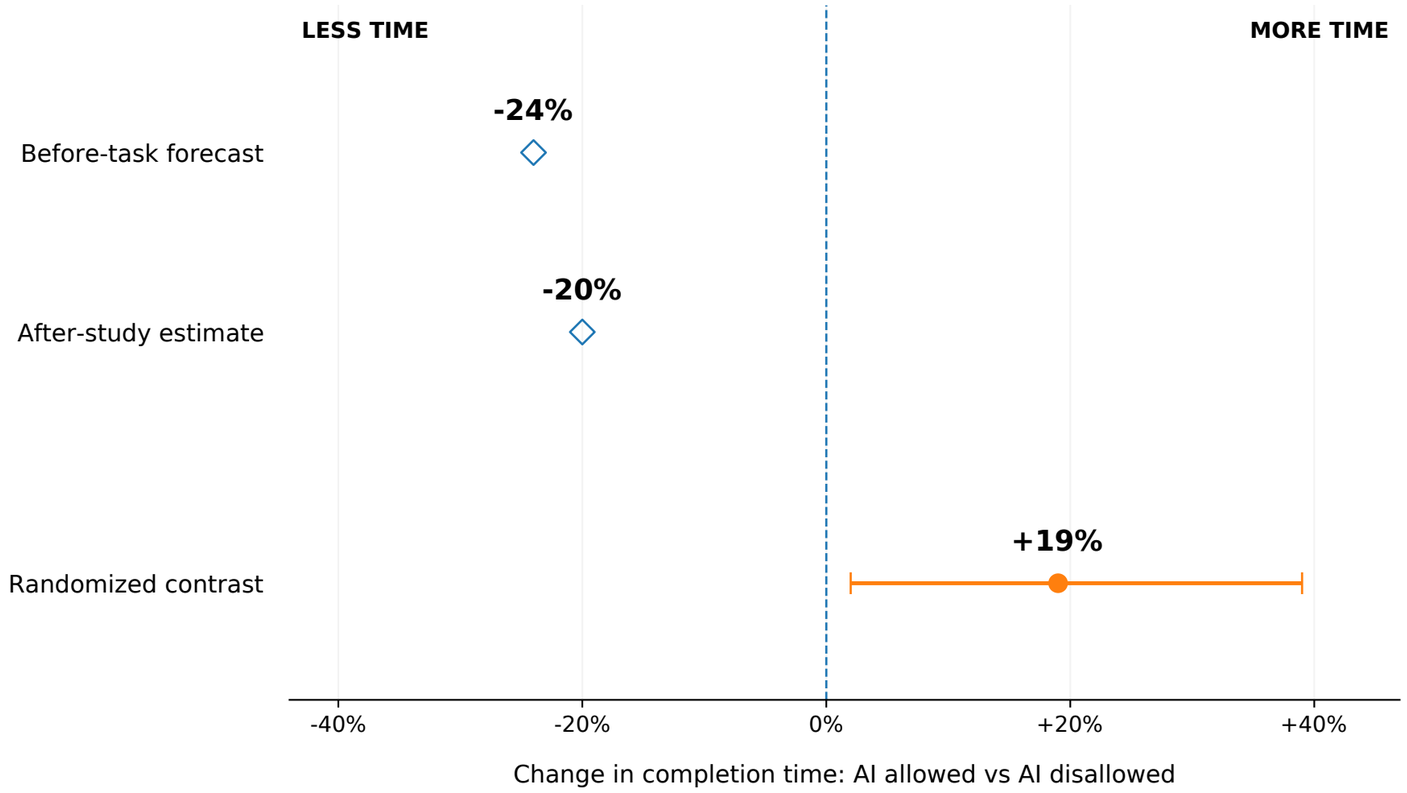
FOR KOSPLORA

Evaluate transferable learning and enjoyment separately. A more engaging world is not automatically a better lesson.

[S7] Wouters et al. (2013), Tables 3–4. Values and 95% intervals reproduced as reported; d is not a percentage.
Subgroup comparisons are observational across studies. They do not isolate the causal effect of randomization.

Feeling faster is not the same as being faster

One task-randomized developer study separated expectations, recollection and the measured time difference.



SETTING

16 developers.
246 real tasks.

Experienced maintainers; February–June 2025 tools. Self-reported time.

LATER EVIDENCE

A 2026 follow-up reports different raw estimates but substantial selection and measurement problems.

FOR THE COMPANION

Ask whether it felt helpful. Also measure time, quality, rework and the outcome the person actually wanted.

[S8] METR (2025), Figure 1 + Appendix D. [S9] February 2026 update confirms the rounded result and interval.
Diamonds = beliefs; intervals omitted. Circle = randomized contrast with reported 95% CI. Historical, not a current benchmark.

Connect the work, and only the work

Illustrative memory and disclosure records. These are proposed controls, not data collected about a real user.

PRIVATE MEMORY

“Worked examples help me learn.”

SOURCE	Explicit statement from the person
PURPOSE	Tutoring examples, not an ability rating
REVIEW	Revisit when the learning goal changes
SHARING	Private until a scoped disclosure is approved

Select + review →

FOR ONE WORKSHOP

“Please offer step-by-step examples.”

SHARE ONLY

This request, its recipient, purpose and expiry.

DO NOT ADD BY DEFAULT

Private chats, unrelated assessments, relationship history or inferred personality traits.

KEEP · EDIT · STOP USING · EXPORT · DELETE

FORGETTING MUST BE OPERATIONAL

Apply a disclosed retention policy to derived summaries, indexes and shared copies, as well as the entry.

Original interface and data-boundary proposal. [S12] informs transaction-level data minimization; [S10] informs access scope.
Deletion, revocation and export require implementation tests. No privacy audit, legal compliance or completed system is claimed.

A booking is not everyone’s commitment

Illustrative shared-workshop request. Different people authorize different parts of the action.

REQUESTER

Approves the booking

Cost, time, cancellation terms and delegated scope.

VENUE

Confirms the space

Availability, access conditions and its own requirements.

INVITEE

Chooses to attend

An invitation is not consent to a task, expense or attendance.

ROOM: CONFIRMED

Submit once; check the venue’s actual response.

ATTENDANCE: PENDING

Do not mark the invitee committed.

Separate records: booking • invitation • accepted commitment • changes • cancellation

RECHECK BEFORE ACTING

Expired scope or material changes mean pause and ask again before acting on an old approval.

Original protocol example. [S10] informs distinct identities, bounded delegation and action checks, but not the social policy.
Statuses shown are hypothetical. No reservation, invitation or message has been created.

Keep the world’s rules outside the narrator

A proposed learning-world architecture separates inspectable simulation state from conversational assistance.

INSPECTABLE MODEL

Player experiment

Choose an action or change a parameter.

→

Simulation core

Apply explicit rules, limits and approximations.

→

State + event log

Record the model version and what changed.

snapshot

MODEL CARD

Show what the world leaves out.

Assumptions

Version + seed

Known limits

Validation checks

Unresolved questions

COMPANION

Explain, ask, offer a hint and cite the model rule.

No silent edits to physics, evidence, scores or action permissions.

OUTSIDE THE GAME: a new task • delayed recall • accessibility supports • independent review

TWO DIFFERENT EVALUATIONS

An agent succeeding inside Minecraft does not establish that the world teaches human learners.

Original architecture proposal. [S6] describes agent learning in Minecraft; [S7] supplies separate evidence about human learning.
A simulation score is not a building approval, physical-safety test or real-world transfer result.

Add physical capability only when it earns its place

Possible forms, not a timetable. Choose the least complex device that meets the person’s need.

FORM	WHAT IT ADDS	EVIDENCE REQUIRED	FALLBACK
Digital	Chosen text, voice and bounded digital actions.	Task benefit, privacy, permission and recovery tests.	Ordinary tools or a non-AI route.
Desk object	Physical presence; optional sensing with visible controls.	Value beyond a phone; guest privacy and access needs.	Remain digital or remove sensing.
Mobile body	Movement within an approved, limited space.	Hazard analysis; localization, stopping and recovery.	Remain stationary; supervised use only.
Task manipulator	One defined handling task with approved objects.	Task-specific validation; failure modes and independent limits.	Remove the task; use human or specialized help.

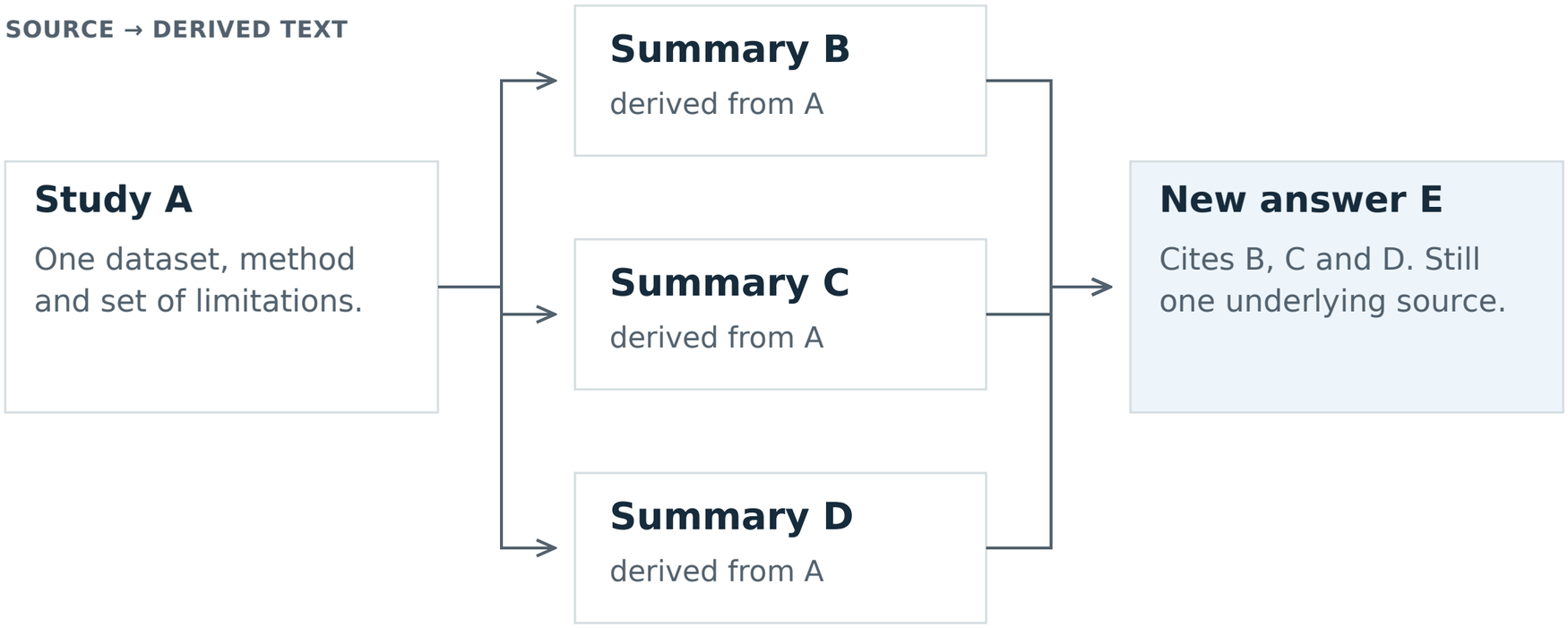
PRESENCE IS NOT PERMISSION

A familiar character does not acquire authority to record guests, enter rooms, touch people or control equipment.

Original staged-capability proposal. [S5] distinguishes model behavior from guaranteed safety-rated stopping.
Deployment needs qualified, context-specific engineering and review. No date, readiness level or safety certification is inferred.

Three summaries can still be one source

Illustrative lineage for AI Searches. Repetition is not independent corroboration.



A DIFFERENT CHECK

Seek independent evidence.

Inspect actual data and methods. A new URL, or another model, does not establish independence.

REUSE GATE: applicability • date • setting • model / method • corrections • remaining uncertainty

PRESERVE THE DEPENDENCY

A correction to A should flag downstream summaries for review, not silently leave the mistake in circulation.

[S11] W3C PROV-O informs derivation records. The shown graph is an original constructed example, not measured network data.
Provenance makes an origin inspectable. It does not prove truth; [S12] makes a related limit explicit for credentials.

Measure the benefit, the burden and the exit

Evaluation sheet for the proposed shared-room pilot. No measurements or completed trials are claimed.

CAPABILITY — NOT MEASURED

Can people explain a new trade-off?

Use a novel task and delayed check; retain accessibility supports.

If not: revise the instruction.

USEFUL WORK — NOT MEASURED

Did the group create something usable?

Seek independent review and evidence of actual use, beyond a completed plan.

If not: simplify the workflow.

BURDEN — NOT MEASURED

Who paid in time, money and effort?

Record rework, facilitation, maintenance and why people declined.

If excessive: narrow the scope.

FREEDOM TO LEAVE — NOT MEASURED

Can people carry on without the system?

Test export, permission revocation and a realistic handoff.

If not: repair the dependency.

DECIDE BEFORE RESULTS ARRIVE

Supportive evidence → repeat elsewhere. Unclear evidence → revise the test. Harm → pause and remedy.

Original outcome and decision framework. “Not measured” is a missing-data state, not zero and not a negative result.
Predefine meaningful effects, methods and pause conditions with participants; do not invent thresholds after seeing the data.

What backs this extension

Two empirical charts, seven proposed system diagrams, and this source register. Different claims need different evidence.

S7 / Wouters et al. · 2013

Serious games meta-analysis

Tables 3–4; historical mixed-design evidence.

S8 / METR · 2025

Developer productivity experiment

Figure 1 + Appendix D; task-randomized study.

S9 / METR · 2026

Follow-up and selection limits

24 February update; not pooled with the first trial.

S10 / NIST · 2026

Agent identity and bounded authority

27 August guidance; not an implementation audit.

S11 / W3C · 2013

PROV-O: provenance relationships

Derivation is visible; correctness remains separate.

S12 / W3C · 2025

Verifiable Credentials v2.0

Authenticity is not truth; minimize requested data.

S5 / Google DeepMind · documentation

Layered physical safety

Developer-reported safeguards; not certification.

S6 / Wang et al. · 2023

Voyager in Minecraft

Agent research, not a trial of human learning.

READING RULE

A source-informed diagram is still a proposal. Only the plotted study results are represented as observed evidence.

Charts: values and uncertainty kept in CSVs with exact source locations. No estimates read by pixel position.
Full links, text alternatives, qualifications and editable SVGs are in the browser edition and source package.