

LIVE WEBINAR · 60 MINUTES

The ARAD Framework

AI-Responsive Assignment Design
for Human Learning

ARAD 2.0

4 modes · 6 principles · 8 dimensions · GOALS



Tim Mousel

Founder, Evolve AI Institute · Creator of the ARAD framework

evolveaiinstitute.com/arad

THE QUESTION FOR THIS HOUR

Does your assignment **still measure** what it claims to measure when a student uses AI?

If your honest answer is "I'm uncertain," you are in good company.



Which assignment in your course feels most vulnerable?

1 Essays or research papers

2 Discussion posts

3 Reflections

4 Problem sets

5 Projects or presentations

6 Exams or quizzes

7 Unsure

VOTE HERE

evolveaiinstitute.com/poll

CODE

SYJZB9

Or type the number in chat. Pick the one you think about most.

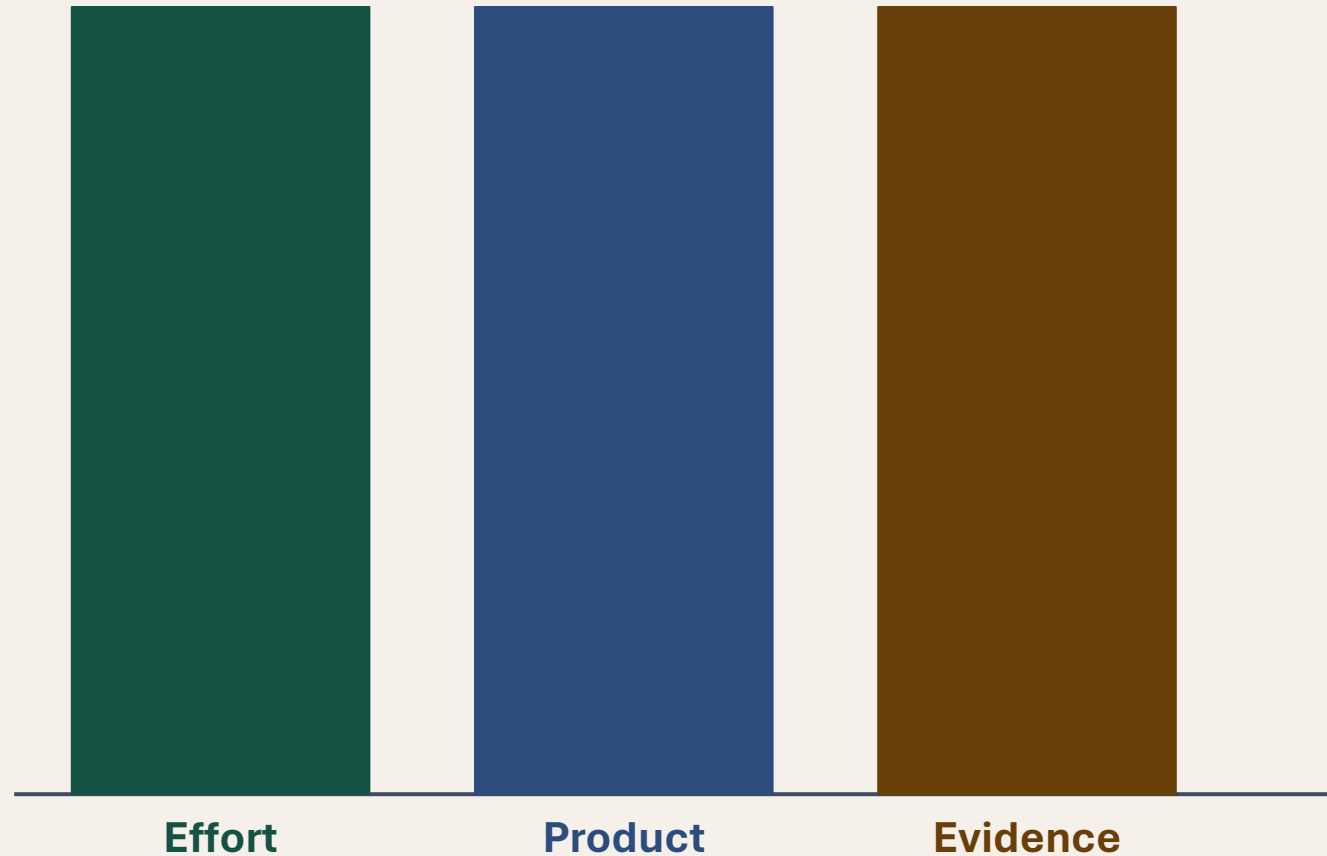
WHY THIS MATTERS NOW

94%

of AI-generated exam submissions went undetected by markers in a blind university study.

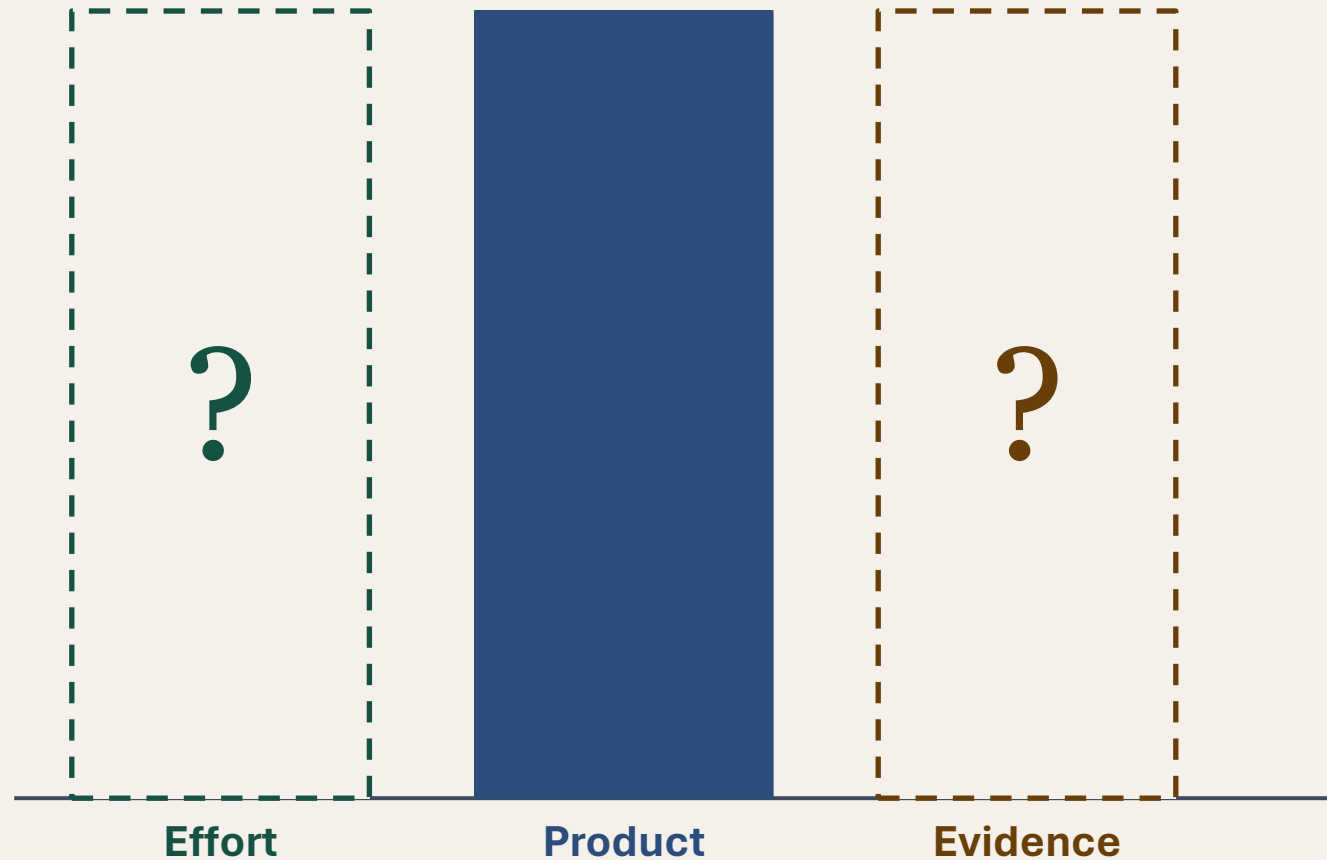
+1/2 grade boundary higher than real student work, on average

For generations, a polished product implied the thinking behind it



Effort, product, and evidence traveled together, so grading the product was a reasonable way to infer the learning.

Now the product can arrive without the thinking



INVISIBLE IN THE PRODUCT ALONE

- Which claims the student tested
- What they rejected, and why
- Where their understanding changed
- Whether they could do it again alone

The design question: what can this assignment still prove?

Retire one question. Ask five better ones.

RETIRE

~~Did the student use AI?~~

Often unanswerable. Even a confident answer leaves the design unchanged.

- | | | |
|---|--|------------------|
| 1 | What must the student learn? | OBJECTIVE |
| 2 | What evidence would demonstrate that learning? | EVIDENCE |
| 3 | What role, if any, should AI play? | AI ROLE |
| 4 | What human thinking must remain visible? | VISIBLE THINKING |
| 5 | How should the assignment and the rubric change? | INSTRUMENT |

ARAD 2.0 CENTRAL QUESTION

Does this preserve the cognitive agency students need?

Four components. One assignment at a time.

4

Modes

decide the role of AI

6

Principles

filter every design

8

Dimensions

evaluate the evidence

5

Steps of GOALS

assemble the redesign

New since ARAD 1.0 (2023): binary becomes a four-mode continuum · process evidence required · verification explicit · equity built in · transfer checkpoint for every major outcome

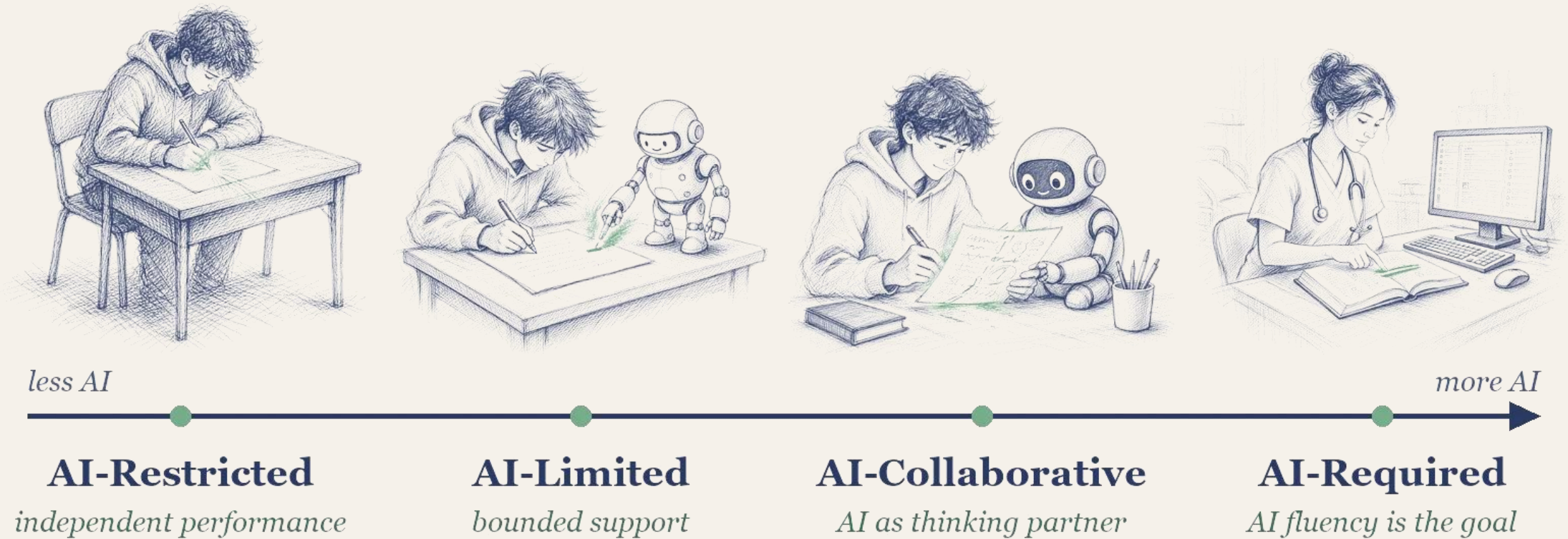
4

PART 1

Four modes

How much AI belongs in this assignment?
The learning goal decides.

Four modes. Equal standing. The goal decides.



Position shows how cognitive work is distributed. Rigor comes from fit with the goal. **Use a mix across a course.**

1 AI-Restricted

The human act is the point

AI is excluded because the assignment measures the student's independent cognitive performance.

WHO DOES THE COGNITIVE WORK



■ Student thinking

Restricted	Limited	Collaborative	Required
------------	---------	---------------	----------

FITS LEARNING GOALS LIKE

Foundational mastery · skill under authentic conditions · independent reasoning · transfer checkpoints

SAY TO STUDENTS

"This task is done without AI because the reasoning is the skill I am assessing, and I need to see yours."

REQUIRED SAFEGUARDS

- A stated pedagogical reason
- Accessible alternatives for students with disabilities
- Fair across language backgrounds and test anxiety

2

AI-Limited

Bounded support, human core

AI may handle bounded support tasks that reduce extraneous load. The intellectual core stays with the student.

WHO DOES THE COGNITIVE WORK



✓ PERMITTED

- Grammar, spelling, formatting
- Brainstorming, then moving past it
- Translation support
- Clarifying the instructions

✗ OFF-LIMITS

- Generating arguments or conclusions
- Writing substantial portions
- Doing the core research or evaluation

SAY TO STUDENTS

"AI may help with grammar, formatting, and translation. The argument and the evidence are yours, and you'll tell me what you used."

REQUIRED SAFEGUARDS

- Boundary statement with the reason
- Brief AI-use disclosure
- Rubric centered on independent thinking

3 AI-Collaborative

The student directs, then judges

AI is used openly inside the learning process. The student directs the thinking; AI contributes material the student evaluates and integrates.

WHO DOES THE COGNITIVE WORK



Restricted Limited Collaborative Required

FITS LEARNING GOALS LIKE

Critiquing arguments and sources · synthesizing perspectives · comparative analysis · iterative revision

SAY TO STUDENTS

"Your grade rests on the quality of your judgment about what the tool produced."

REQUIRED SAFEGUARDS

- Full transparency: prompt logs, transcripts, rationales
- Structured verification with a concrete deliverable
- Rubric that scores the student's judgment
- At least one human performance checkpoint

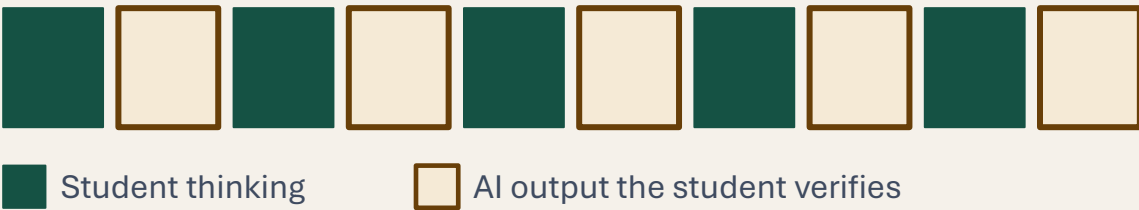
4

AI-Required

Skilled use is the competency

AI use is mandatory because skilled human-AI collaboration is the professional competency being taught.

WHO DOES THE COGNITIVE WORK



Restricted Limited Collaborative **Required**

FITS LEARNING GOALS LIKE

Clinical AI scribes · CTE diagnostic workflows · AI-assisted CAD and code · paralegal research · data analysis

SAY TO STUDENTS

"Find at least one error in the AI's output and correct it against an authoritative source, cited by page or section."

REQUIRED SAFEGUARDS

- Scenario authenticity
- Discipline-standard verification
- One cited correction per task
- Oral or observed defense
- A Transfer Gap Check

Choose the mode: four assignments, four decisions

A

NURSING

Students chart three simulated encounters using the clinical AI scribe deployed at the partner hospital, then submit a corrected chart.

B

STATISTICS

Students select the right hypothesis test for eight described studies and justify each choice. The instructor wants students to hold the decision rules themselves.

C

COMPOSITION

Students write a 900-word essay developing an original argument about a text read in class. Several students are multilingual writers.

D

MARKETING

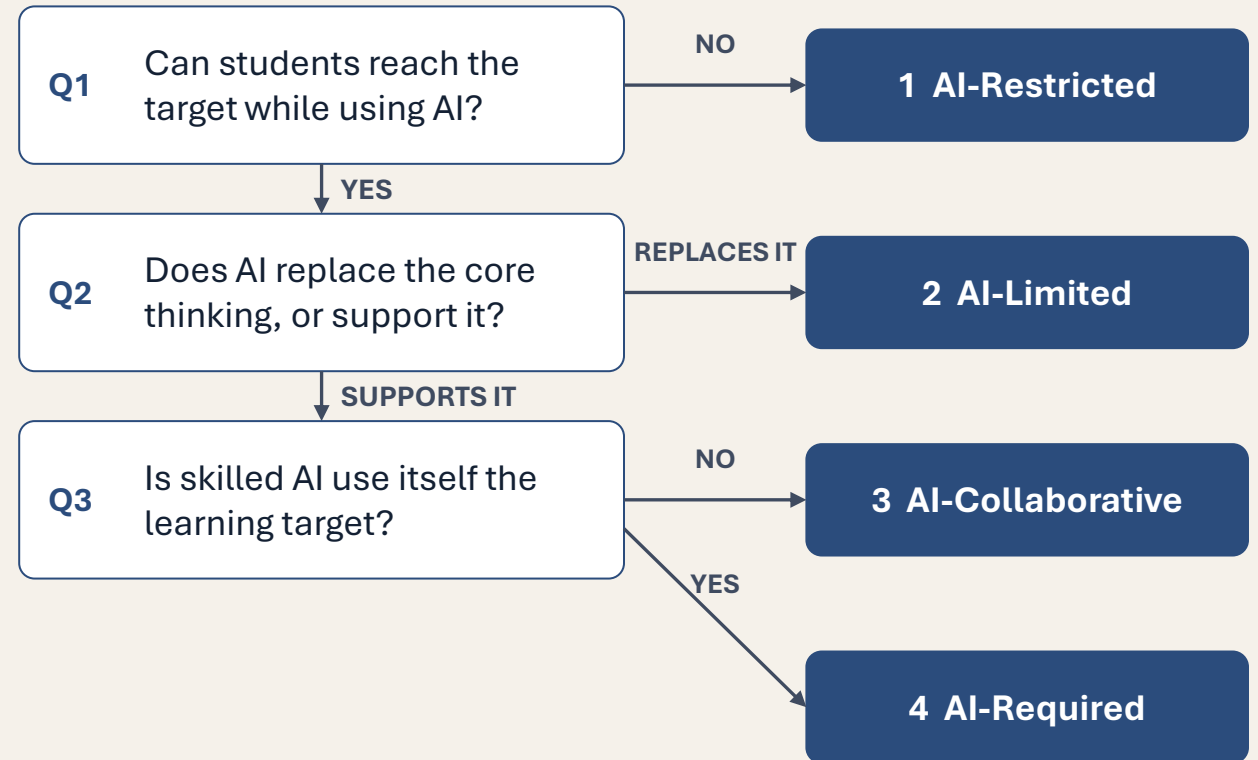
Students produce a competitive analysis of a regional company and recommend a market-entry strategy.

Type four answers in chat, like **A4 B1 C2 D3** 1 Restricted · 2 Limited · 3 Collaborative · 4 Required

The objective decides the mode

A	Nursing	4 AI-Required
B	Statistics	1 AI-Restricted
C	Composition	2 AI-Limited
D	Marketing	Depends on the objective

D: applying Porter's Five Forces to compare strategies fits Mode 3. Building original analysis from local primary research fits Mode 2.



Always: one AI-Restricted Transfer Gap Check per major outcome.

6

PART 2

Six principles

Design commitments that double as quality filters. Fail any one, and that part gets redesigned.

Run every assignment through all six filters

1 Cognitive Agency

Who is doing the thinking at each stage?

2 Learning Integrity

Does it still measure what it claims if a student uses AI?

3 Process Visibility

Can I see how the student got there?

4 Evaluative Judgment

Must the student judge AI output against course standards?

5 Productive Struggle

Has AI removed the difficulty that makes learning stick?

6 Equity and Access

Can every student do this fairly, independent of AI detection?

1

Cognitive Agency

The student stays the primary thinker, decision-maker, and judge. AI may support, extend, or challenge; the student drives the cognitive work.



THE FILTER

Who is doing the thinking at each stage?

THE TWO-MINUTE STAGE AUDIT

Choose the topic	Student	
Gather the evidence	Student	
Build the argument	Student	← carries the objective
Produce the draft	Shared	
Revise	Student	

Example Teacher education: the student chooses which misconceptions to target and how to sequence the lesson. AI may draft the materials list.

2

Learning Integrity

The assignment must still measure the intended outcome when AI is part of the process. If AI lets students bypass the target skill, the assessment loses construct validity.



THE FILTER

Does it still measure what it claims if a student uses AI?

THE 20-MINUTE FACULTY PREREQUISITE

PROMPT

- 1 Run it through two or three current models, the way a motivated student would.

SCORE

- 2 Grade the output on your own rubric. If it earns an 87, record the 87.

INTERPRET

- 3 AI at Proficient or above where human capability is required? The rubric measures the wrong thing.



The graded AI response, annotated, becomes teaching material for a Verification Challenge.

3

Process Visibility

Every assignment requires evidence of how the student thought, in addition to what they produced. Visible thinking is the primary way to tell learning from AI-generated output.



THE FILTER

Can I see how the student got there?

SEVEN ARTIFACTS, ORDERED BY WHAT THEY COST YOU TO READ

Baseline memo before AI contact		LOW
Annotated bibliography with rationale		LOW
Prompt log and transcript		LOW
Marked-up AI draft		MEDIUM
Revision notes		MEDIUM
Narrated screen recording		HIGH
Observed in-class artifact		HIGH

One well-chosen artifact usually transforms an assignment. Exemplary Process Evidence needs at least one observed-setting artifact.

4

Evaluative Judgment

Students critically assess, test, and improve AI output, treating it as a draft to challenge. Verification is the substantive exercise of disciplinary judgment.



THE FILTER

Must the student judge AI output against course standards?

BEGINNING LEVEL

"I checked the AI's facts."

"Some parts seemed off."

A reflection written after the fact

EXEMPLARY LEVEL

Corrects one claim, citing a source by page or section

Names the standard the output violated

Correction submitted with the work

Why page numbers? A specific citation sends the student into the text itself, where the grounding lives.

5

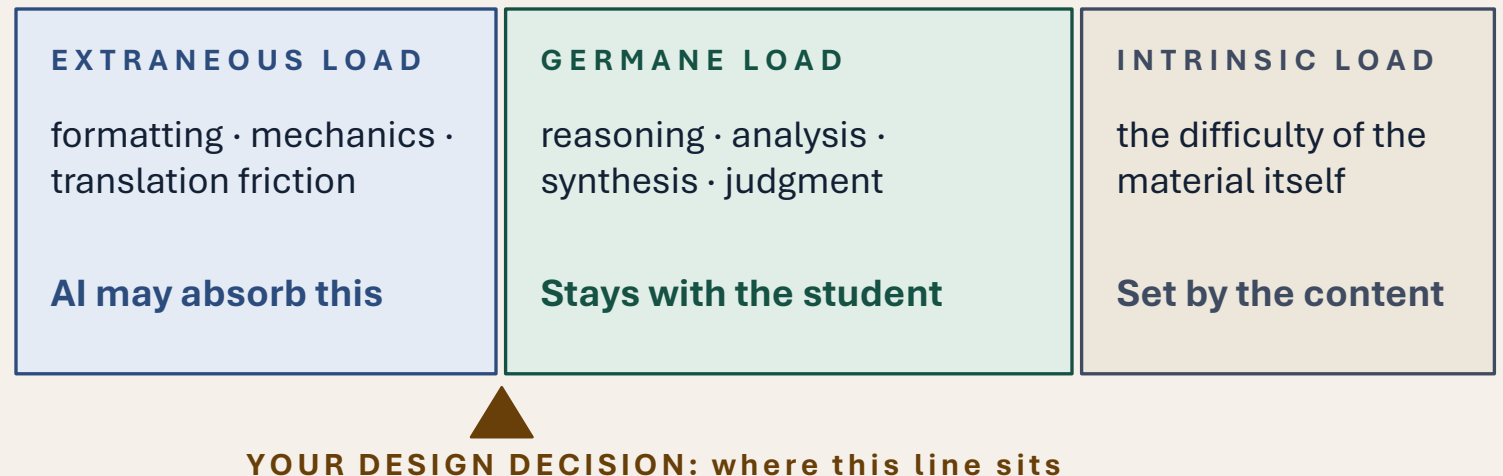
Productive Struggle

Preserve the effort that produces durable learning. AI may reduce extraneous difficulty, and the germane difficulty that builds understanding stays.



THE FILTER

Has AI removed the difficulty that makes learning stick?



Example Automotive: AI may summarize the service bulletin. The student still works the diagnostic tree on the vehicle.

6

Equity and Access

Fair, accessible, and achievable for every student regardless of tool access, digital literacy, language background, or disability. ARAD rejects AI detection as a grading or misconduct tool.



THE FILTER

Can every student do this fairly, independent of AI detection?

FOUR OBLIGATIONS THIS PRINCIPLE CREATES

1 Provide access, or run the AI work in class

Free and paid tiers produce different output.

2 Teach the tool before you grade its use

A 30 to 45 minute AI literacy primer first.

3 Offer an opt-out pathway

Same grade ceiling, workload, and documentation. In the syllabus before the term.

4 Keep detection out of grading

Detectors are unreliable and biased against non-native English writers (Liang et al., 2023).

8

PART 3

Eight dimensions

A rubric that scores the student's thinking, built so the top level is hard to fake.

Eight dimensions. Choose the three or four your goal needs.

1

Cognitive Agency

Does the student own the intellectual work?

2

Evaluative Judgment

Do they assess AI output against disciplinary standards?

3

NEW

Process Evidence

Is the thinking visible and documented?

4

NEW

Metacognitive Depth

Do they show awareness of their own thinking?

5

NEW

Prompt Quality

How well do they direct the AI?

6

Verification Quality

Did they test, challenge, and improve the output?

7

Disciplinary Judgment

Do they apply course knowledge and standards?

8

Transfer

Can they perform independently in a new context?

Levels Beginning → Developing → Proficient → Exemplary Exemplary always demands evidence that is hard to fake.

Scoring a familiar assignment before students see it

"Write a 1,200-word research paper about a course-related topic using five scholarly sources."

Cognitive Agency	B D P E	Developing	Topic is the student's; who decided what stays invisible
Evaluative Judgment	B D P E	Beginning	The task omits claim testing
Process Evidence	B D P E	Beginning	One submission, one moment, zero observed artifacts
Verification Quality	B D P E	Beginning	Five sources counted, zero tested
Transfer	B D P E	Beginning	Independent checkpoint is missing

THREE SMALL REPAIRS

1

Add a baseline artifact

2

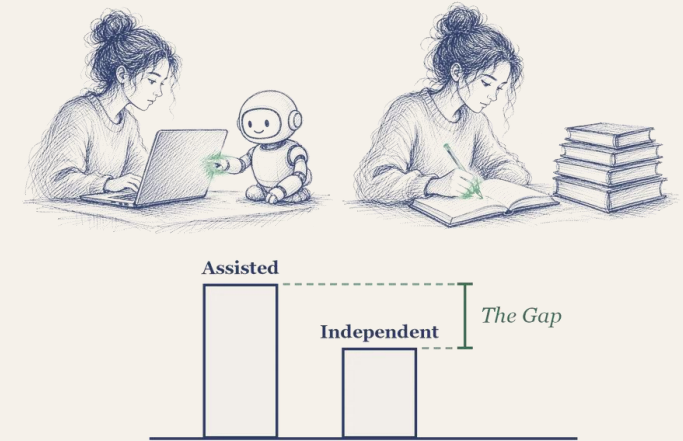
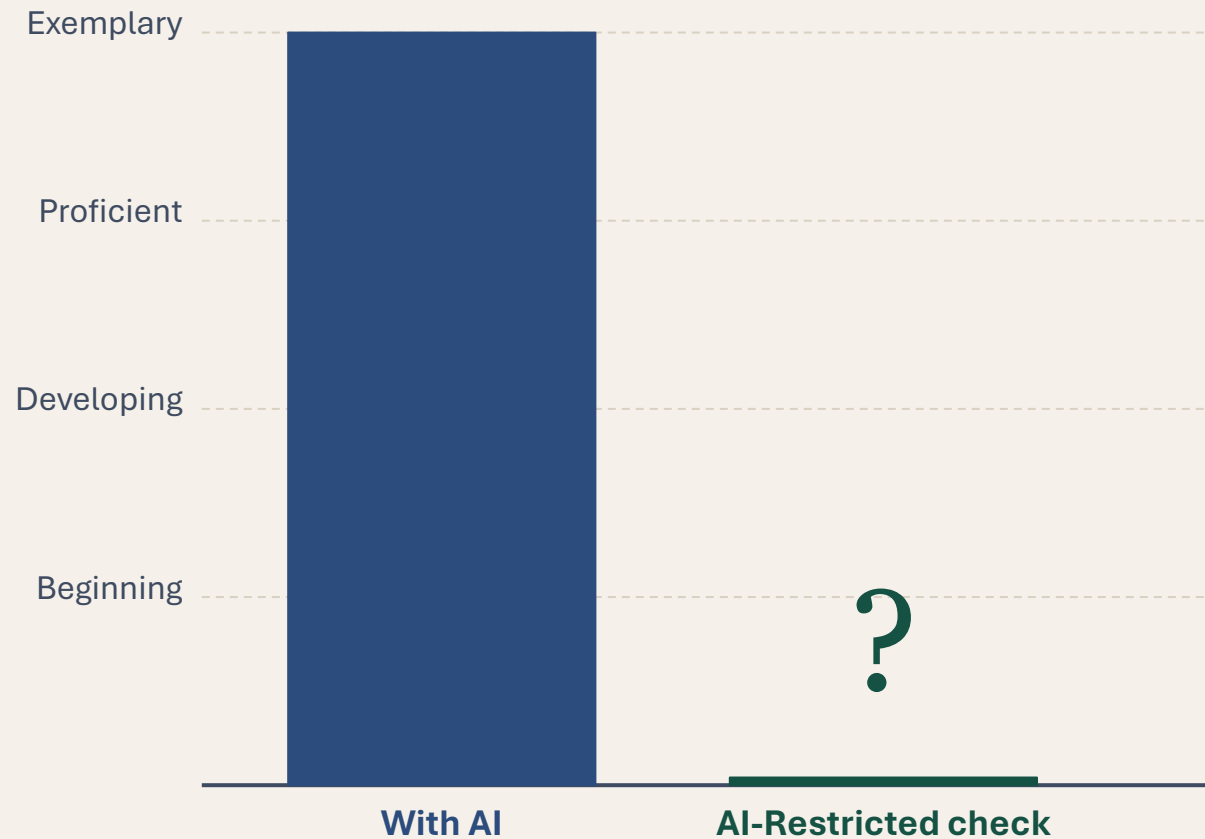
Require one cited correction

3

Add an independent checkpoint

The weakness is missing evidence, and the repairs are small.

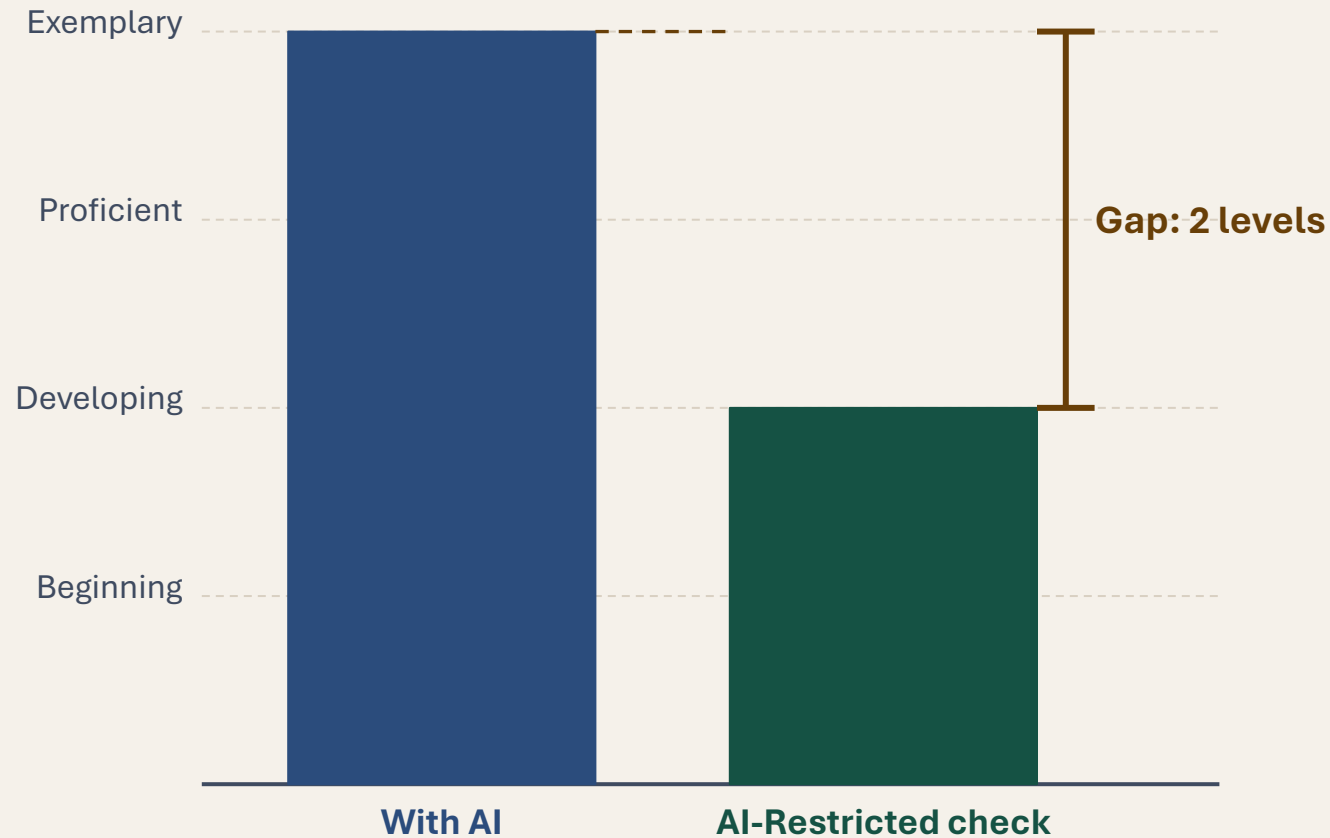
Same rubric dimensions, with AI and without it



THE RULE

Every major course outcome is assessed at least once under AI-Restricted conditions, on the same dimensions as the assisted task.

The gap turns a grade into a finding



GAP OF ONE LEVEL OR LESS

The learning transferred. The student holds the capability independently.

GAP GREATER THAN ONE LEVEL

The learning is tool-dependent. Redesign the upstream task toward a more restricted mode, or raise the weight of the restricted checkpoint.

GOALS

PART 4

The GOALS process

Five steps. Each one produces something you can hold.

Five steps, five tangible outputs

Source: goal-first sequencing follows Backward Design (Wiggins & McTighe, 2005)

G

Goal

Define the human learning target

OUTPUT

a one-sentence target statement

O

Openness

Choose the right AI mode

OUTPUT

a mode plus a rationale shared with students

A

Architecture

Design for visible thinking

OUTPUT

a phased task sequence

L

Look for Evidence

Align the rubric

OUTPUT

3 or 4 dimensions with descriptors

S

Sustain

Scaffold, pilot, and maintain

OUTPUT

a literacy plan and a review date

↻ **Annual review sends S back to G.** One assignment, start to finish: 15 to 30 minutes once you've done it before.

Target first. Mode second.

G · GOAL: NAME THE HUMAN CAPABILITY

A FORMAT POSING AS A GOAL

"Write an essay about climate policy."

A HUMAN LEARNING TARGET

"Evaluate competing policy arguments using evidence and construct an original, defensible position."

Ask: what part of this thinking must remain human for learning to occur?

O · OPENNESS: FOUR QUESTIONS, IN ORDER

1 Does AI support or replace the target learning?

2 Which specific cognitive acts must remain human?

3 Would AI remove a desirable difficulty?

4 Can all students access and use the tools?

Write the rationale down and share it. Students follow a stated reason.

Architecture: two phases, four patterns

PHASE A · FORMATIVE

Learning phase. AI often more open.

Planning document · structured AI interaction with logs · peer review · instructor checkpoint · reflection

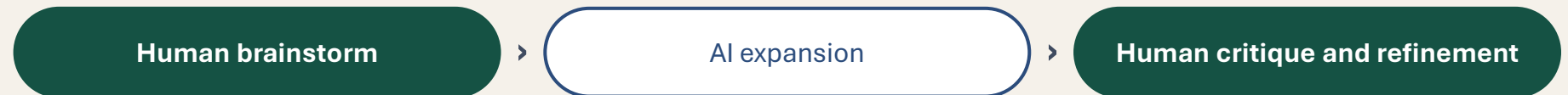
PHASE B · SUMMATIVE

Demonstration phase. Verification raised.

Final deliverable with process documentation · verification protocol · human performance checkpoint · transfer task

FOUR PATTERNS · FILLED = STUDENT STEP · OUTLINED = AI STEP

AI Sandwich



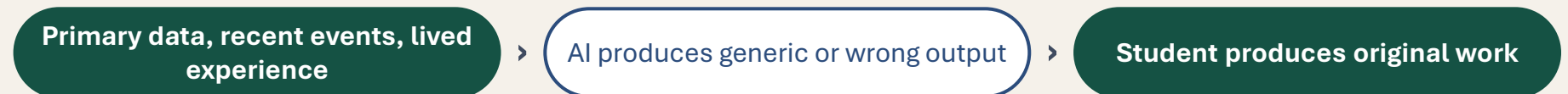
Verification Challenge



Comparative Analysis



Generative Challenge



Align the rubric. Then keep the design alive.

L · LOOK FOR EVIDENCE

- 1 Select 3 or 4 of the eight dimensions
- 2 Write descriptors for at least two levels
- 3 Add a Transfer Gap Check if this maps to a major outcome

A redesigned assignment graded on the old rubric reverts to the old assignment.

S · SUSTAIN: THREE LINKED JOBS

SCAFFOLD

A 30 to 45 minute AI literacy primer before the graded task.

STUDY

Collect one form of data: score distribution, process artifacts, or targeted student feedback.

SUSTAIN

Re-run the assignment through current models every year. Keep a capability watchlist.

An assignment left unvisited drifts back into an ARAD 1.0 assignment.

Before and after: the same paper, redesigned

BEFORE

"Write a 1,200-word research paper about a course-related topic using five scholarly sources."

Graded on thesis, evidence, organization, and mechanics. One upload in week 8.

Evaluative Judgment	Beginning
Process Evidence	Beginning
Verification Quality	Beginning
Transfer	Beginning



AFTER · AI-COLLABORATIVE · VERIFICATION CHALLENGE

- 1 **Claim memo, in class, 20 min, before any AI contact**
- 2 Source set: five sources, two sentences on why each
- 3 AI episode: three competing claims; full transcript submitted
- 4 **Paper, 1,000 words: adjudicate the claims; correct one, cited by page**
- 5 Reasoning note: what changed, what remains open
- 6 Oral checkpoint, 3 minutes
- 7 **AI-Restricted transfer task, in class, 20 minutes**

Three changes did the work: baseline artifact · cited correction · independent checkpoint. Student time: 9 hours (was 10).

Your turn. One assignment, seven questions.

- 1 What is the most important learning outcome?**
- 2 What could a student outsource to AI while bypassing that learning?
- 3 Which ARAD mode fits?**
- 4 What evidence of human reasoning, performance, or judgment do you need?
- 5 What task feature should be added, removed, or changed?**
- 6 Which rubric criterion should be revised?
- 7 What must students be told about acceptable AI use?

Short on time? Answer 1, 3, and 5.

0:40

Pick one assignment that matters

4:00

Silent work. Notes are fine.

1:20

In chat: discipline · mode · one change

[Worksheet: evolveaiinstitute.com/arad](https://evolveaiinstitute.com/arad)

Three signals that a design needs to tighten

1

Near-identical, AI-shaped output

Submissions cluster around the same phrasing and structure.

RESPONSE

Require an observed-setting artifact. Consider moving the core task to AI-Limited.

2

Transfer gap above one level

Restricted scores fall more than a level below assisted scores.

RESPONSE

Demand more baseline thinking before AI contact, or raise the restricted checkpoint's weight.

3

Top-clustered assisted scores

Tight at the top with AI, spread out on the restricted check.

RESPONSE

Audit Cognitive Agency, Evaluative Judgment, and Prompt Quality for their anti-gaming anchors.

For program chairs: every graduate should have two independent Transfer Gap Checks per major outcome, sequenced across levels and scored by different instructors.

THE MISTAKE THAT CUTS BOTH WAYS

Assuming AI integration
is automatically
innovative.

Assuming AI restriction is
automatically rigorous.

Both skip the objective.

If you feel confident about the AI decision before you've written the learning target, slow down.

Start with one assignment. Monday morning, 30 minutes.

20 min

Run the assignment through two current models and score the output honestly.

5 min

Write the one-sentence human learning target.

5 min

Name the one artifact that would make the thinking visible.

Then pilot it for one term, collect one form of data, and revise for the next offering.

DECIDE BEFORE LAUNCH

- Tool access
- Integrity policy alignment
- Accessibility of new artifacts
- Privacy: what goes into tools
- Opt-out pathway
- Multilingual support

Two assistants walk the process with you

CUSTOM GPT BY TIM MOUSEL

AI-Responsive Assignment Design (ARAD) Framework

Bring one assignment. It walks the GOALS steps with you: goal, mode, architecture, rubric, and sustainment.

CUSTOM GPT

Create AI-Resistant or AI-Dependent Assignments

Built for the two ends of the spectrum: designs that exclude AI, and designs built around required AI use.



evolveaiinstitute.com/arad

Slides, worksheet, both assistants, and
the full article

Bring your sprint notes. The assistant takes it from there.

THE CENTRAL MESSAGE

Begin with what students must learn and what evidence will make that learning **visible**.

The AI decision follows from those two answers.

ONE THING THIS WEEK

Run one assignment through a current model, score it on your own rubric, and write the number down.



Questions

Your question, or one line in chat: what will you change first?

4 modes

Restricted · Limited · Collaborative · Required

6 principles

Agency · Integrity · Visibility · Judgment · Struggle · Equity

8 dimensions

Agency · Judgment · Process · Metacognition · Prompting · Verification · Discipline · Transfer

GOALS

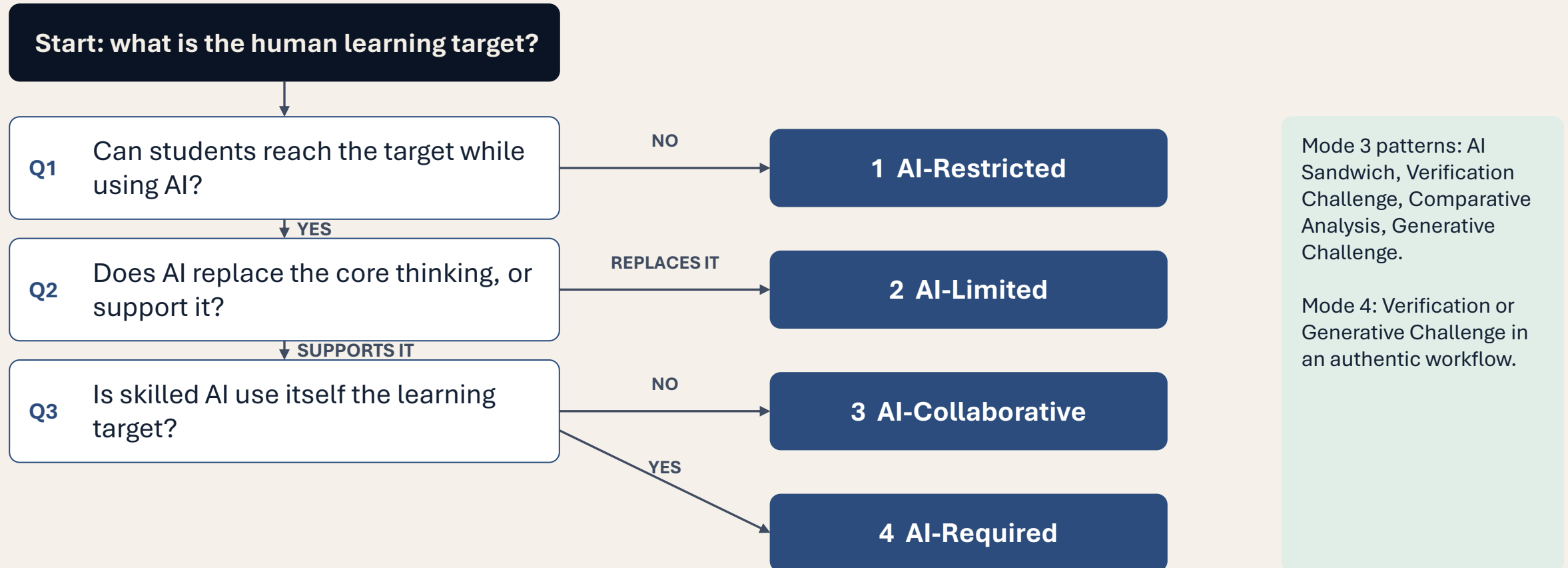
Goal · Openness · Architecture · Look for Evidence · Sustain



evolveaiinstitute.com/arad

Tim Mousel · Evolve AI Institute

Quick-reference decision flowchart



Always: every major outcome gets at least one AI-Restricted Transfer Gap Check, on the same dimensions as its assisted task.

The eight dimensions: what Exemplary and Beginning look like

Dimension	Exemplary	Beginning
1 Cognitive Agency	Clearly drives every key decision; AI stays subordinate to their judgment.	Accepted AI output with minimal engagement.
2 Evaluative Judgment	Names specific strengths and weaknesses; corrects at least one AI claim citing a course reading by page or section.	Accepts AI output uncritically or offers generic corrections.
3 Process Evidence	At least one observed-setting artifact alongside transcripts and written work.	Only the final product exists.
4 Metacognitive Depth	Identifies specific learning moments and residual uncertainty; explains prompt choices with evidence.	Generic, post-hoc, formulaic reflection.
5 Prompt Quality	Specific, iterative prompts with deliberate strategy that evolve with the output.	One vague prompt; output accepted wholesale.
6 Verification Quality	Systematic verification against multiple sources, testing logic and framing.	Output accepted as-is.
7 Disciplinary Judgment	Integrates course concepts throughout with deep understanding.	Generic reasoning without disciplinary grounding.
8 Transfer	Applies the learning in an unfamiliar context without AI support.	Shows the learning only with AI support.

Twelve predictable mistakes

DESIGN MISTAKES

01

One AI policy applied to every assignment

02

Choosing the mode before clarifying the objective

03

Assessing polished output without assessing reasoning

04

Redesigning every assignment at once

INSTRUCTION MISTAKES

05

Requiring disclosure without defining meaningful disclosure

06

Prohibiting AI without giving the pedagogical reason

07

Permitting AI without teaching verification

08

Adding process documentation until it becomes busywork

JUDGMENT MISTAKES

09

Treating AI detection as definitive evidence

10

Revising the assignment and leaving the rubric unchanged

11

Assuming AI integration is automatically more innovative

12

Assuming AI restriction is automatically more rigorous

References

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 100779. <https://doi.org/10.1016/j.patter.2023.100779>

Mousel, T. (2026). ARAD 2.0: AI-Responsive Assignment Design for Human Learning. Evolve AI Institute. <https://evolveaiinstitute.com/edai/blog/arad-2-ai-responsive-assignment-design.php>

Scarfe, P., Watcham, K., Clarke, A., & Roesch, E. (2024). A real-world test of artificial intelligence infiltration of a university examinations system: A "Turing Test" case study. *PLOS ONE*, 19(6), e0305354. <https://doi.org/10.1371/journal.pone.0305354>

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296. <https://doi.org/10.1023/A:1022193728205>

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). Association for Supervision and Curriculum Development.