

SPECIAL REPORT • AUGUST 2026

The Defensive AI Trap

Why competitive evaluation systems force cognitive offloading — and how to redesign them.

DOC TYPE	DATA ANCHOR	FOCUS	CYCLE
Special Report	Brown University	Incentive Design	Aug 2026

Executive Summary

Generative AI adoption in learning and professional environments is no longer driven primarily by laziness or a desire to shortcut effort. It is driven by game-theoretic panic. When evaluation systems reward inflated output and enforce strict grade curves, individuals adopt AI defensively — to avoid taking a self-imposed penalty against peers who use it.



Empirical Anchor: The Brown University Data Case

In July 2026, data from a Brown University course revealed a stark performance divergence: on a take-home midterm exam, the class average reached 96%. When brought into a supervised room for the final exam, the average plummeted to 48%.

96% → 48%

Take-home vs. in-person score crash, same cohort, same material.

88% of Students

Say they worry generative AI is actively eroding their cognitive capacity and critical thinking skills — an internal Brown survey found.

The paradox: despite knowing the long-term cognitive cost, students still used AI overwhelmingly on take-home assignments. Awareness of the harm didn't change the behavior — because the incentive driving it was never really about the harm.

Academic Foundation: The Social Dynamics of AI

Three peer-reviewed papers explain why individuals act against their own stated self-interest:

Defensive AI Adoption [Bursztyn et al., NBER Working Paper No. w34488, 2025](#)

Demonstrates that social comparison drives defensive AI adoption. Parents and students adopt AI tools not out of intrinsic belief in their efficacy, but out of fear that non-adoption places them at an insurmountable disadvantage.

The Comparative Trap [Zhang & Lee, IJHCI, 2026](#)

Proves that high social comparison orientation directly fuels problematic generative AI use, overriding personal ethical concerns.

Post-Secondary AI Anxiety [Farooqi et al., arXiv:2601.10468, 2026](#)

Documents how career and job anxiety forces CS students into aggressive tool reliance to keep pace with algorithmic expectations.

Macro Mechanics: Drivers of Defensive AI

Two structural conditions turn individually rational AI use into a system-wide trap:

01

Zero-Sum Curves & Caps

When institutions cap top marks — Harvard-style A-caps — or grade on tight curves, students compete head-to-head. Submitting unassisted work starts to feel like academic self-sabotage.

02

Output-Only Evaluation

Evaluating finished static artifacts — essays, code snippets, cover letters — rewards whoever leverages the highest model capacity, not whoever understands the material best.

Strategic Solution: Process Over Output

Surveillance tools, honor codes, and AI detectors fail because they try to police a game-theoretic incentive trap with software. The only sustainable fix is structural:

- **Invert the assessment loop:** grade the active synthesis and oral or interactive defense of ideas, not the submitted text file.
- **Process-visible environments:** evaluate live, collaborative problem-solving where student reasoning is transparent in real time.
- **De-escalate curve pressure:** replace artificial zero-sum ranking with mastery-based criteria that reward genuine skill acquisition.

Actionable Implementation Blueprint

A prompt for faculty to stress-test their own assignments before students find the shortcut first:

COURSE REDESIGN PROMPT • COPY IN FULL

Open Claude, ChatGPT, or Google Gemini. Input your current assignment prompt and ask:

"Generate 3 ways a student could complete this assignment in under 5 minutes using LLMs without reading the source material. Then suggest 2 structural modifications that require real-time collaborative reasoning or live defense."

About Nowcollab & Primary Sources

Nowcollab builds collaborative learning and active assessment frameworks for higher education faculty, course directors, and academic leaders.

Inside Higher Ed (July 2026)

Bursztyn et al., NBER w34488 (2025)

Zhang & Lee, IJHCI (2026)

Farooqi et al., arXiv:2601.10468 (2026)