

PRESERVING CRITICAL THINKING

In the Face of AI

SLIDES & CITATIONS

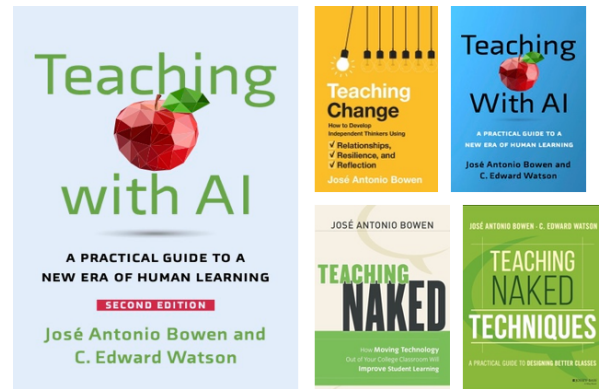
José Antonio Bowen

Get the [2nd Edition of Teaching with AI](#).

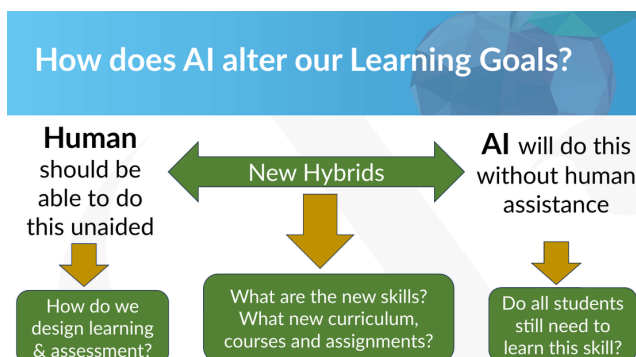
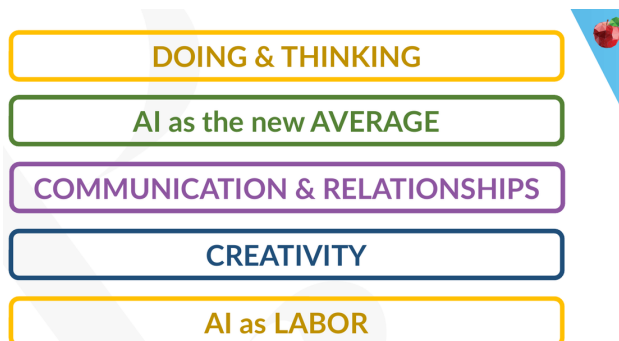
Discount code HTAI25 = 30% off at JHUP.

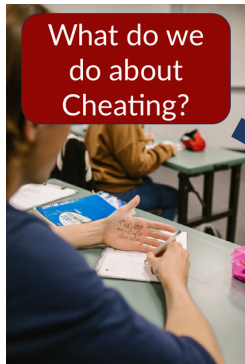
WebSite: <https://WeTeachWithAI.com/>

► [Workshop Prompts & Materials](#)



Curious
AND
Skeptical





RESEARCH FINDINGS



Marken, S. (2026, April 2). AI is routine for college students, despite campus limits. Gallup Education.

Oakley, B., Johnston, M., Chen, K.-Z., Jung, E., & Sejnowski, T. J. (2025). The memory paradox: Why our brains need knowledge in an age of AI. SSRN [preprint — final venue not yet verified].

The Conversation. (2026). College students are writing with AI, but a pilot study finds they're not simply letting it write for them. [small pilot] <https://theconversation.com/college-students-are-writing-with-ai-but-a-pilot-study-finds-theyre-not-simply-letting-it-write-for-them-276856>

Chirikov, I., Smirnov, I., & Kizilcec, R. F. (2026). Generative AI use and misuse call for assessment reform in higher education. Science, 392(6800), 818–820.

<https://doi.org/10.1126/science.aec5115> [95,513 students at 20

US public research universities; substantial heterogeneity in GenAI use and misuse across disciplines and student groups; misuse estimated by indirect questioning, not self-report]

Stephenson, R., & Armstrong, C. (2026). Student generative AI survey 2026 (HEPI Report 199). Higher Education Policy Institute.

<https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/> [1,054 UK undergraduates; 12% put AI-generated text directly into assessed work, up from 8% in 2025 and 3% in 2024]

Angehr, E., Bloem, M., Howell, J., & Radford, A. W. (2026, February). College faculty perceptions of generative artificial intelligence in higher education. College Board. <https://research.collegeboard.org/media/pdf/ai-research-brief-3-vf.pdf> [3,000+ US faculty, fielded June–August 2025; 74% say students use AI to write essays, 77% use AI themselves, 45% negative sentiment — faculty perceptions, not measured student behaviour]

Cruces, G., Fernández Mejjide, D., Galiani, S., Gálvez, R. H., & Lombardi, M. (2026). Does generative AI narrow education-based productivity gaps? Evidence from a randomized experiment (NBER Working Paper No. 34851). National Bureau of Economic Research. <https://doi.org/10.3386/w34851> [n=1,174; education gap 0.548 SD falls to 0.139 SD with AI, but returns to 0.200 SD in a follow-up task without it]

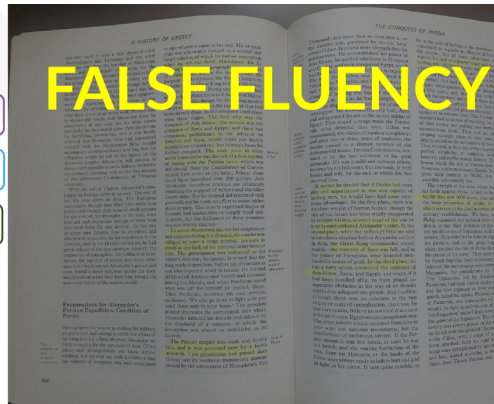
AI Synthesis

Less time engaging

Shallower knowledge

"Metacognitive Laziness"

Stadler et al. (2024)
Melumad & Yun (2025)
Kreijkes et al. (2025)
Fan et al. (2025);
Lee et al. (2025);
Kreijkes et al. (2025);
Kosmyrna et al. (2025, preprint);
Gerlich (2025, corrected)



of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530. [RCT, 117 students: essay scores up, no gain in knowledge or transfer] <https://doi.org/10.1111/bjet.13544>

Lee, H.-P., et al. (2025). The impact of generative AI on critical thinking. CHI '25. [confidence in AI ↔ less critical thinking] <https://doi.org/10.1145/3706598.3713778>

Kosmyrna, N., et al. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing. arXiv [preprint; small sample — cite as a cultural phenomenon, not as evidence]. <https://arxiv.org/abs/2506.08872>

Gerlich, M. (2025). AI tools in society. *Societies*, 15(1), 6. [self-report perception data, 666 participants; formally corrected September 2025] <https://doi.org/10.3390/soc15010006>

Melumad, S., & Yun, J. H. (2025). Experimental evidence of the effects of large language models versus web search on depth of learning. *PNAS Nexus*, 4(10), pgaf316. <https://doi.org/10.1093/pnasnexus/pgaf316>

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>

Kreijkes, P., et al. (2025). Effects of LLM use and note-taking on reading comprehension and memory: A randomised experiment in secondary schools. SSRN [preprint — final venue not yet verified].

Fan, Y., et al. (2025). Beware of metacognitive laziness: Effects

When students use AI without guidance

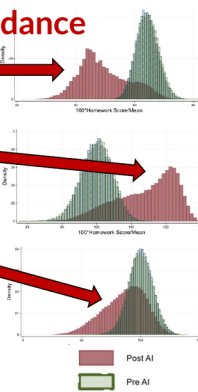
↓ 30% Time on Homework

↑ 18% Homework Grades

↓ 20% Exam Scores

Largest Learning Loss

- Social Science, STEM and Languages
 - Younger, High-Achieving & Boys.
 - Shorter Homework Completion Time
- (26,811 Chinese Students in grades 7-12)



Strömberg, D., Lei, V., & Wu, Y. (2026). The generative AI learning penalty: Evidence from Chinese secondary education (CEPR Discussion Paper No. 21577). Centre for Economic Policy Research. [working paper]

<https://cepr.org/publications/dp21577>

Weeks, J. O., Voshaar, J., Plate, B. J., & Zimmermann, J. (2025). Generative AI usage and exam performance. ECIS 2025 Proceedings. <https://arxiv.org/abs/2404.19699>

Rismanchian, S., et al. (2026). Faster completion, less learning: Generative AI reduced study time on math problems and the knowledge they build. arXiv [preprint].

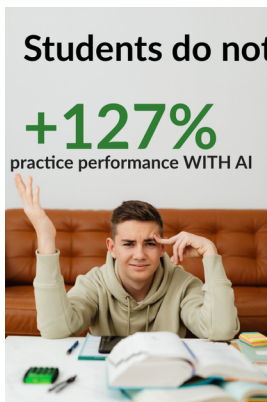
<https://arxiv.org/abs/2605.21629>

Students do not perceive the damage

+127%
practice performance WITH AI

-17%
on the exam (v. no AI control)

AI disrupts self-assessment



Bastani et al. (2025) PNAS, RCT n=1,000 · Weeks et al. (2025)
Fernandes et al. (2026) *Computers in Human Behavior*, N=246 +
replication N=452

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *PNAS*, 122, e2422633122. [RCT, n ≈ 1,000]

<https://doi.org/10.1073/pnas.2422633122>

Barcaui, A. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. *Social Sciences & Humanities Open*, 12, 102287. [RCT, n = 120; surprise test 45 days later: 57.5% vs 68.5%, d = 0.68 — single site, note attrition]

<https://doi.org/10.1016/j.ssaho.2025.102287>

Fernandes, D., et al. (2026). AI makes you smarter but none the wiser: The disconnect between performance and

metacognition. *Computers in Human Behavior*, 175, 108779. [N = 246; replication N = 452; higher AI literacy predicted worse metacognitive accuracy] <https://doi.org/10.1016/j.chb.2025.108779>

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. <https://doi.org/10.1037/0022-3514.77.6.1121>

Lodge, J. M., & Loble, L. (2026). Artificial intelligence, cognitive offloading and implications for education. University of Technology Sydney. [the “performance paradox”] <https://doi.org/10.71741/4pyxmbnjq.31302475.v2>

AI TOOLS that MAKE STUDENTS DO THE WORK

Step-by-step Guidance

Answers Withheld

Mastery Pacing

Hallucination & Expertise
Guardrails

Prompts that Push



More Engagement
More Learning
Less Time

Kestin et al. (2025); Deng et al (2024); Chung et al. (2026);
De Simone et al. (2025) World Bank; Bastani et al. (2025)

Kestin, G., Miller, K., Klales, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning. *Scientific Reports*, 15, 17458. [Harvard RCT, N = 194]

<https://doi.org/10.1038/s41598-025-97652-6>

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. [69 studies; higher-order gains largely from subjective student reports]

<https://doi.org/10.1016/j.compedu.2024.105224>

Chung, A. T.-H., Zhang, B., Kung, L.-C., Bastani, H., & Bastani, O. (2026). Effective personalized AI tutors via LLM-guided reinforcement learning. *SSRN* [working paper; +0.15 SD

unassisted exam gain $\approx$ 6–9 months of schooling] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6423358

De Simone, M. E., et al. (2025). From chalkboards to chatbots: Evaluating the impact of generative AI on learning outcomes in Nigeria.

World Bank [working paper; 800 students; girls gained most] <http://documents.worldbank.org/curated/en/099548105192529324>

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>

Sourati, Z., Karimi-Malekabadi, F., Ozcan, M., McDaniel, C., Ziabari, A., Trager, J., Tak, A. N., Chen, M., Morstatter, F., & Dehghani, M.

(2026). The shrinking landscape of linguistic diversity in the age of large language models. *Nature Human Behaviour*. [880,000+ texts, 7 datasets; LLM polishing cut writing-complexity variance 21-50% while keeping content, and amplified dominant patterns]

<https://doi.org/10.1038/s41562-026-02550-0>

McQuaid, F., Skafida, V., McQuillan, R., & Lee, B. (2026). Screen use and emotional, behavioral, and social development in children. *JAMA Network Open*, 9(8), e2626804. [Scottish birth cohort 2009–2020; no association at ages 5 or 10; at 15, 1-2 h/day social media OR 0.57 = fewer difficulties, 5+ h/day gaming OR 2.61] <https://doi.org/10.1001/jamanetworkopen.2026.26804>

Macro Buddy (GPT Custom Bot)

Saharnaz Babaei Balderlou

University of Wisconsin–La Crosse



130 Students:

Exams in person, on paper, proctored, no AI

Control (no AI, work alone) -

Groups (no AI) -

Alone with AI Tutor Bot +12%

Groups with AI Tutor Bot +14%



Babaei-Balderlou, S., & Shakya, S. (2026). Can structured AI tutoring and study groups improve learning outcomes? Evidence from a randomized controlled trial. *The Journal of Economic Education*, 1–18. <https://doi.org/10.1080/00220485.2026.2695361>

Saharnaz Babaei-Balderlou, University of Wisconsin–La Crosse.

► [Macro Buddy \(GPT custom bot\)](#)

Babaei-Balderlou, S., & Shakya, S. (2026). Can structured AI tutoring and study groups improve learning outcomes? Evidence from a randomized controlled trial. *The Journal of Economic Education*, 1–18. [RCT, 4 sections, 130 students; the first randomized evidence of AI-enhanced learning in economics education]

<https://doi.org/10.1080/00220485.2026.2695361>

Babaei-Balderlou, S., & Shakya, S. (2025, July 25). The invisible hand of Gen-AI: Can AI-enhanced study groups improve

learning outcomes? *SSRN*. [working-paper version, the one cited on the slide] <https://doi.org/10.2139/ssrn.5696622>



WeTeachWithAI.com

► <https://weteachwithai.com/designing-learning-with-ai/>

Li, C., Cui, H., & Hagedorn, L. S. (2026). The cognitive impact of ChatGPT in higher education: A systematic review of critical and creative thinking outcomes. *Computers and Education: Artificial Intelligence*, 10, Article 100571. [PRISMA review of 67 empirical studies, 2022–2025 — the strongest synthesis to date] <https://doi.org/10.1016/j.caeai.2026.100571>

Wang, J., & Fan, W. (2025). RETRACTED (22 April 2026): The effect of ChatGPT on students' learning performance... *Humanities and Social Sciences Communications*, 12, 621. [RETRACTED — never cite as evidence] <https://doi.org/10.1057/s41599-025-04787-y>

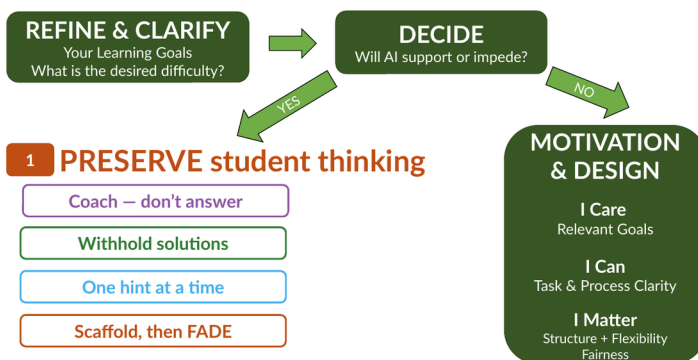
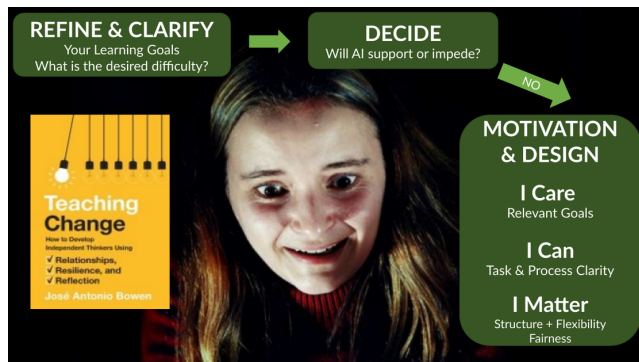
Bartoš, F., Martinková, P., & Wagenmakers, E. (2025, May 12). Adjusting for publication bias reveals no evidence for the

effect of ChatGPT on students' learning performance. [preprint reanalysis behind the retraction]

https://doi.org/10.31234/osf.io/8vs32_v1

LIMITS. The strongest evidence of harm comes from secondary students outside the United States (China, Turkey, Germany, the U.K.); the strongest evidence of benefit comes from elite, high-structure settings such as Harvard. The typical university classroom sits in the understudied middle, where age, selectivity and design are confounded. Most of these studies measure exam scores; we still lack long-term evidence about durable, transferable learning.

DESIGN PRINCIPLES



Bowen, J. A. (2021). *Teaching change: How to develop independent thinkers using relationships, resilience, and reflection*. Johns Hopkins University Press. [I care / I can / I matter]

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364.

<https://doi.org/10.1007/BF00138871>

Kestin, G., et al. (2025). AI tutoring outperforms in-class active learning. *Scientific Reports*, 15, 17458.

<https://doi.org/10.1038/s41598-025-97652-6>

Chung, A. T.-H., et al. (2026). Effective personalized AI tutors via LLM-guided reinforcement learning. SSRN [working paper].

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6423358

CUSTOM BOTS = HIDDEN PROMPT + KNOWLEDGE BASE (slide 17)

A custom bot is just two things: a hidden system prompt that fixes the behavior (coach, don't answer; one hint at a time; never give the solution) and a knowledge base that fixes the content.

All of the custom tools and simulations in this session were stress-tested on the best models but set to run and interact with students using only free models.

- [Examples, templates and instructions for building simulations and custom bots](#)



Li, C.-J., Hwang, G.-J., Chang, C.-Y., & Su, H.-C. (2025). Generative AI-supported progressive prompting for professional training. *British Journal of Educational Technology*, 56(6), 2550–2572. [quasi-experimental, n = 62] <https://doi.org/10.1111/bjet.13594>

Kalyuga, S., & Plass, J. L. (2025). *Rethinking cognitive load theory*. Oxford University Press.

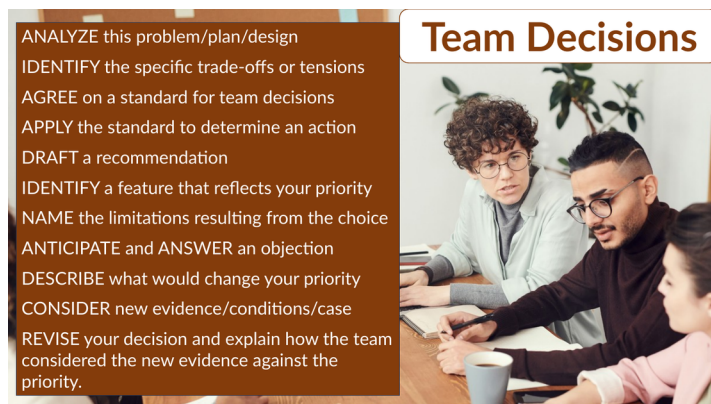
<https://doi.org/10.1093/9780190078539.001.0001>

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38(1), 23–31. https://doi.org/10.1207/S15326985EP3801_4

Vaccaro, M., Almaatouq, A., & Malone, T. W. (2024). When combinations of humans and AI are useful. *Nature Human Behaviour*, 8, 2293–2303. [106 experiments: worse on decision tasks, synergy on creation tasks]

<https://doi.org/10.1038/s41562-024-02024-1>

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>



Example Assignment

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In *Psychology and the Real World* (2nd ed., pp. 59–68). Worth.

de Bruin, A. B. H., et al. (2023). Worth the effort: The Start and Stick to Desirable Difficulties (S2D2) framework. *Educational Psychology Review*, 35(2), 41.

<https://doi.org/10.1007/s10648-023-09766-w>

Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424.

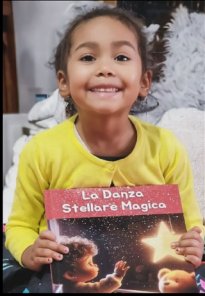
<https://doi.org/10.1080/07370000802212669>



GENERIC or CUSTOMIZED?



Touchnote



Scarlett Panda

Greeting Cards
Bedtime Stories
Birthday Songs

Treatment Plans
Instructions
Textbooks
Assignments
Exams



- [Touchnote](#)
- [Scarlett Panda](#)

Increase Relevance & Customization



Sweller, J. (1988); Clark, R., Nguyen, F., & Sweller, J. (2006); Paas, F., Renkl, A., & Sweller, J. (2003)

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
https://doi.org/10.1207/s15516709cog1202_4

Clark, R., Nguyen, F., & Sweller, J. (2006). Efficiency in learning: Evidence-based guidelines to manage cognitive load. Pfeiffer.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4.
https://doi.org/10.1207/S15326985EP3801_1

3

AMPLIFY to raise the bar

Fast, private, formative feedback at scale

What more can you do WITH AI?

AI tempts students into less thinking: Require more if it.

Meyer, J., Jansen, T., Schiller, R., Liebenow, L., Steinbach, M., Horbach, A., & Fleckenstein, J. (2023). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*.
<https://doi.org/10.1016/j.caeai.2023.100199>

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
<https://doi.org/10.3102/003465430298487>

AI as Idea Generator



Spicy Mango Jerky
Matcha Green Tea Popcorn
Gluten-Free Brazilian Cheese Bread Mix
Hibiscus Ginger Lemonade
Rose Petal Jam

- Iterate & Improve
- Cost & Price
- Packaging & Size
- Market Test
- Supply Chain & Margin
- Propose the Best 10



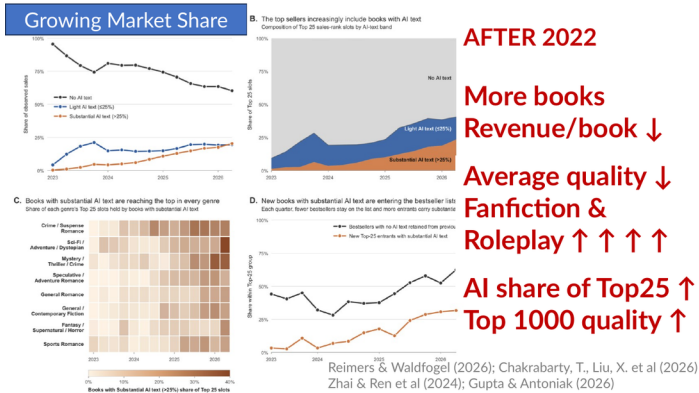
Urban, M., Děchtěrenko, F., Lukavský, J., Hein, V., Švácha, F., Brom, C., & Urban, K. (2023). ChatGPT improves creative problem-solving performance in university students: An experimental study. [preprint]
<https://doi.org/10.31234/osf.io/9z2tc>

Slide 24: VIDEO

► [Video embedded in this slide \(YouTube\): https://www.youtube.com/watch?v=odtw7E1KPMY](https://www.youtube.com/watch?v=odtw7E1KPMY)

Reimers, I., & Waldfogel, J. (2026). AI and the quantity and quality of creative products: Have LLMs boosted creation of valuable books? NBER Working Paper 34777. [working paper] <https://doi.org/10.3386/w34777>

Chakrabarty, T., Liu, X., et al. (2026, July). Generative AI floods and dilutes the market for books. arXiv [preprint]. <https://arxiv.org/pdf/2607.20349>



The fanfiction and roleplay line comes from the WildChat corpus: in more than 500,000 consented, English-language ChatGPT conversations, over a third involve generating fiction, especially fanfiction and erotica which generate what the authors call a "solipsistic reader-writer," "who both generates and consumes fiction within a closed conversational loop."

Gupta, N., Antoniak, M., & Walsh, M. (2026, June 23). AI fiction in the wild. arXiv [preprint; provisionally forthcoming in MFS: Modern Fiction Studies]. <https://arxiv.org/abs/2606.22748>

Zhao, W., Ren, X., Hessel, J., Cardie, C., Choi, Y., & Deng, Y.

(2024). WildChat: 1M ChatGPT interaction logs in the wild. ICLR 2024. [consented corpus of 1M conversations; the dataset behind Gupta et al.] <https://arxiv.org/abs/2405.01470>



<https://doi.org/10.3389/feduc.2025.1697554>

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning. Psychological Science, 17(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>

Singh, A., Guan, Z., & Rieh, S. Y. (2025). Enhancing critical thinking in generative AI search with metacognitive prompts. Proceedings of ASIS&T, 62(1), 672–684. <https://doi.org/10.1002/pa2.1287>

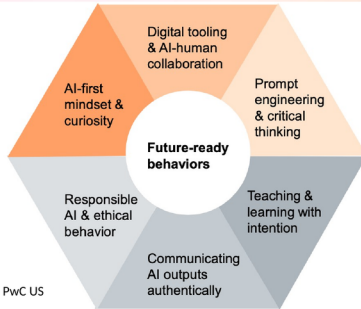
Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. British Journal of Educational Technology, 56(5), 1842–1863. <https://doi.org/10.1111/bjet.13599>

Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. Learning and Instruction, 95, 102011. <https://doi.org/10.1016/j.learninstruc.2024.102011>

Tomisu, H., Ueda, J., & Yamanaka, T. (2025). The cognitive mirror: A framework for AI-powered metacognition and self-regulated learning. Frontiers in Education, 10, 1697554. [framework paper, not empirical]



Future-ready behaviors



© 2025 PwC US

AI Usage
**MONITORED
& REQUIRED**

4x/day
MINIMUM

AI use at PwC is monitored and required. PwC US CEO Paul Griggs: anyone who thinks they can “opt out” of AI is “not going to be here that long.” Accenture staff must demonstrate “regular adoption,” with usage tracked, for promotions.

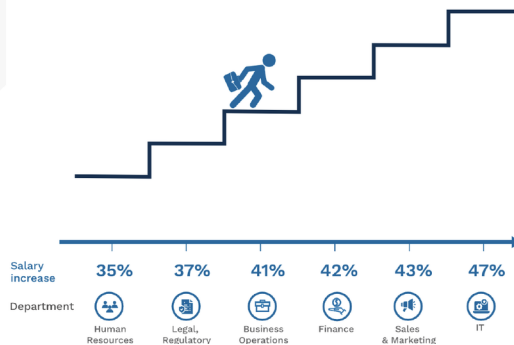
Accounting Today. (2026, March). PwC CEO: You cannot opt out of AI.

The Register. (2026, March 19). PwC AI.

https://www.theregister.com/2026/03/19/pwc_ai/

AI and Salaries

Salary boost to employees acquiring AI skills, by department¹
Weighted average percentage increase in salary indicated by employers



Access Partnership & Amazon Web Services. (2023, November). Accelerating AI skills: Preparing the workforce for jobs of the future. [employer and worker survey commissioned by AWS; 1,340 U.S. organizations, 3,297 workers]

<https://assets.aboutamazon.com/e1/a0/17842ee148e8af9d55d10d75a213/aws-accelerating-ai-skills-us-en.pdf>

Make Thinking Visible

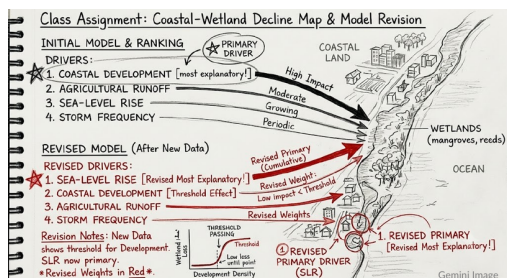
Fewer Artifacts - More Process & Judgement

- Pre-Game** [PRIORITIZE - ANALYZE - RANK - CLARIFY
DECIDE - CHOOSE
- Draft** [WRITE - DEFEND - JUSTIFY - RECOMMEND - EXPLAIN
-- AI or no AI
- Feedback** [READ - DEBATE - DISCUSS - ENCOUNTER - COMPARE
--New evidence, conditions, cases, priorities, perspectives (from AI?)
- After Party** [REVISE - RESPOND - REFLECT
--COMMENT in your original doc (!) what you now
--STRENGTHEN, REMOVE, CHANGE, REFRAME, RESHAPE

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In Handbook of Self-Regulation (pp. 13–39). Academic Press.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In Metacognition: Knowing about Knowing (pp. 185–205). MIT Press.

This framework is heavily influenced by a new book: Sanford, D. R. (2027, forthcoming). Artificial intelligence, authentic voices: AI & assignment design for critical thinking in higher education. Bloomsbury.



Daniel R. Sanford, Director, CTL, Boise State University
Artificial Intelligence, Authentic Voices: AI & Assignment Design for Critical Thinking in Higher Education (2027) Bloomsbury



This Coastal Wetlands example is from his new book. You can learn more at his website. Dan is the Director of the CTL at Boise State University.

► [Daniel R. Sanford — Artificial Intelligence, Authentic Voices](#)

BUILD a map with causal arrows explaining coastal-wetland decline
RANK all four drivers by how much explanatory work each does
MARK the driver you judge most explanatory.
CONSIDER (one week later) more data on the same region.
ADD new drivers and revise the weights on your original map. Use a new color.

Teach Students How to Use the Bot!

45% understanding fundamental concepts and jargon
42% practice questions

"I'm terrible at prompting,
so I don't think I'm getting much out of it"

HALF never logged on

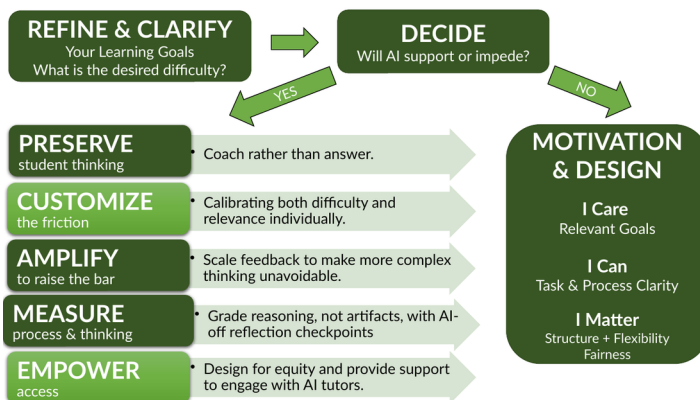
Sax, M. (2025) Harvard Business School Course Tutor Bot 930 First-year accounting students
Robinson et al. (2026) Annenberg/Stanford • 2 RCTs, 355 elementary students

Sax, M. (2025, June 23). Custom AI tutor bots are transforming learning at HBS. Harvard Business Impact. [930 first-year accounting students; only 12% became power users] <https://hbsp.harvard.edu/inspiring-minds/ai-tutor-bots-harvard-business-school>

Robinson, C. D., Gormley, D., Ribeiro, A. T., & Loeb, S. (2026). Access is not enough: Human support improves engagement with AI tutoring (EdWorkingPaper No. 26-1451). Annenberg Institute, Brown University. [working paper; 2 RCTs, 355 elementary students; engagement rose 71–80% but reading achievement did not] <https://doi.org/10.26300/pz7p-p388>
Henderson, M., et al. (2025). Comparing generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. Assessment & Evaluation in Higher Education. [8,642 open-ended responses: teachers more

trusted] <https://doi.org/10.1080/02602938.2025.2502582>

Er, E., et al. (2025). Assessing student perceptions and use of instructor versus AI-generated feedback. British Journal of Educational Technology, 56(3), 1074–1091. <https://doi.org/10.1111/bjet.13558>

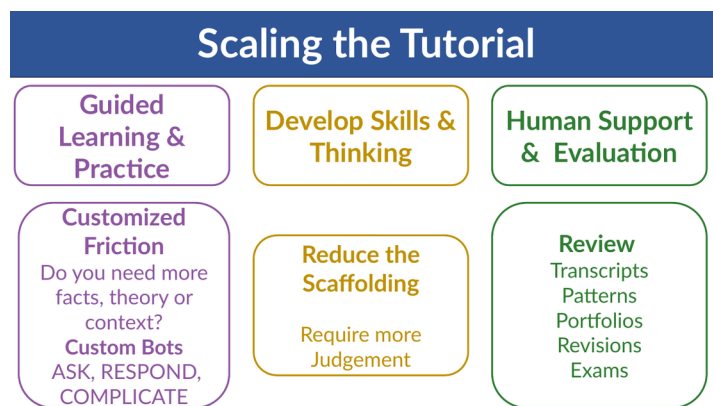


► [Full principles, templates and examples](#)

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demsky, D. (2024). Tutor CoPilot: A human-AI approach for scaling real-time expertise. arXiv [preregistered RCT: 900 tutors, 1,800 K-12 students; +4 points mastery, +9 for the lowest-rated tutors, ~\$20/tutor/year] — principle 8, EQUIP the humans. <https://arxiv.org/abs/2410.03017>

Martin, A. J., et al. (2025). Integrating generative AI and load reduction instruction. Learning and Individual Differences, 121, 102723. <https://doi.org/10.1016/j.lindif.2025.102723>

Dunn, T. L., & Risko, E. F. (2016). Toward a metacognitive account of cognitive offloading. Cognitive Science, 40(5), 1080–1127. <https://doi.org/10.1111/cogs.12273>



Bloom's tutorial advantage was never mostly about content delivery. It was guided practice, immediate correction and a human who noticed. AI scales the first two cheaply; on the current evidence the third still has to be human, and the gains show up only when the tool withholds answers and someone reviews the process. This is the whole ReDesign argument in one frame: the six boxes are the eight principles applied to a single course.

Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13(6), 4–16. <https://doi.org/10.3102/0013189X013006004>

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. Review of Educational Research, 86(1), 42–78. [the 2-sigma claim does not replicate at scale; ITS effects cluster near 0.4–0.6 SD] <https://doi.org/10.3102/0034654315581420>

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demsky, D. (2024). Tutor CoPilot: A human-AI approach for scaling real-time expertise. arXiv [preregistered RCT: 900 tutors, 1,800 K-12 students] <https://arxiv.org/abs/2410.03017>

Oreopoulos, P., & Low, N. (2026). One click away: AI tutoring with Khanmigo in a two-year school experiment (NBER Working Paper No. 35620). National Bureau of Economic Research. [COUNTER-EVIDENCE. Cluster RCT, 18 Tennessee middle schools, 53 grade clusters, two years; not yet peer-reviewed. Khanmigo was configured to coach rather than give answers and still produced only 0.04 SD pooled (0.02

year 1, 0.08 year 2) — no better than Khan Academy practice without AI. 96% tried it, but the median student messaged it on a third of practice days and in 17% of sessions where they made a mistake. Withholding answers is not enough; engagement is the binding constraint.] <https://www.nber.org/papers/w35620>

Jin, Y., Martinez-Maldonado, R., Gašević, D., Han, X., & Yan, L. (2026). Emergent learner agency in implicit human–AI collaboration: How supportive and contrarian AI personas reshape interaction. *Journal of Computer Assisted Learning*, 42(5), e70310. [COUNTER-EVIDENCE. RCT, 224 students randomised into 97 online triads; 10-minute creative task; the AI teammate was undisclosed. A contrarian AI produced the Idea→Challenge→Integration pattern in 38.2% of groups against 13.3% (control) and 12.1% (supportive), but lowered teamwork satisfaction and psychological safety ($H(2)=12.14$, $p=.002$) with no gain in cognitive load or creative performance. Designed friction is not the same as learning. Not pre-registered; five motifs tested and two p-values sit just under .05.] <https://doi.org/10.1002/jcal.70310>

EXAMPLES

Fact-Checking Prompt Contest



1. Contribute one claim to our shared document.
2. Write a prompt that verifies facts and sources.
3. Test it. Improve it. Explain why (comments).
4. Test prompts in pairs.
5. Compare to Mike Caulfield's Deep Background prompt/GPT.
6. Write a critical thinking process for yourself that uses AI.
7. Reflect (bot?): why is critical "doing" also important?



- ▶ [Complete assignment](#)
- ▶ [Deep Background GPT \(fact checks and context\)](#)
- ▶ [Complete SIFT prompt](#)

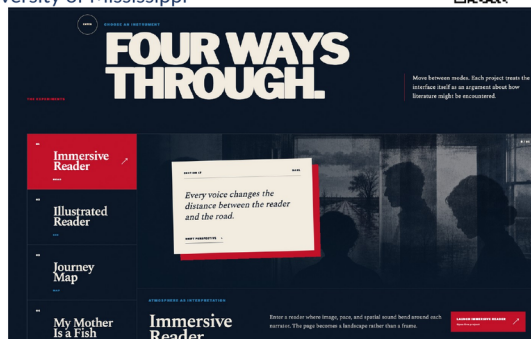
Reading Faulkner's As I Lay Dying

Marc Watkins, University of Mississippi



No embedded chatbots, Faulkner personas, automatic summaries, or buttons that tidy away the novel.

Each project is an encapsulated module designed to help students slow down, explore, annotate, and make an interpretive choice.



Reading Faulkner's As I Lay Dying

(slide 36)

Marc Watkins, University of Mississippi.

- ▶ [As I Lay Dying gallery](#)
- ▶ [Marc Watkins Substack](#)

ART HISTORY FINAL: Teach an AI how to paint like John Constable. Create a new image and clarify what stylistic elements are visible.



Ira Greenburg, Director of the Center of Creative Computation and Professor of Art at SMU

AI Disclosure

- Where did you use AI?
- What did you do without AI?



AI Literacy

- Where did AI make your work better or worse?
- What AI work did you alter?
- What would you have missed without AI?
- How did you know?

Broad Street, 1854

You are Dr John Snow. Find the killer.



On a laptop
josebowen.github.io/
BroadStreet1854/

Three Mile Island, 1979

You can stop the meltdown.



On a laptop
josebowen.github.io/
TMI1979

- [Broad Street, 1854](#)
- [Three Mile Island, 1979](#)

Make Thinking the Assignment

AI
Inputs



Convene a Meeting
Moderate a Debate
Choose a Path
Prioritize a Value
Hold a Symposium
Decide a Strategy

Human
Outputs



Invert the usual order. Let AI produce the raw material — the brief, the data, the opposing positions — and make the human work the part that cannot be delegated: convening, weighing, choosing and defending. The gradable artifact is the judgment, not the prose.

- [Assignment examples](#)

Singh, A., Guan, Z., & Rieh, S. Y. (2025). Enhancing critical thinking in generative AI search with metacognitive prompts.

Proceedings of ASIS&T, 62(1), 672–684. <https://doi.org/10.1002/pa2.1287>

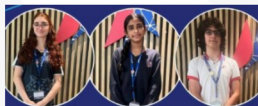
Vaccaro, M., Almaatouq, A., & Malone, T. W. (2024). When combinations of humans and AI are useful. *Nature Human Behaviour*, 8, 2293–2303. [106 experiments: human-AI teams did worse on decision tasks, better on creation tasks — an argument for keeping the decision human] <https://doi.org/10.1038/s41562-024-02024-1>

STUDENT ASSIGNMENTS

AI can do (some) things humans can't

PathoScan Screens for Colorectal Cancer

- Trained on US public gut microbiome data to find anomalies.
- Looks for bacterial imbalances and biomarkers associated with early-stage colon cancer.



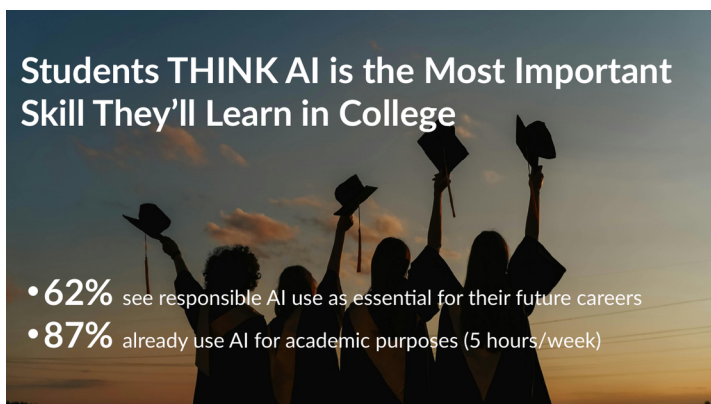
Dubai American Academy ISTE
AI Challenge Winners

► [ISTE AI Innovator Challenge winners](#)



Marken, S. (2026, April 1). College students weigh AI's impact on majors and careers. Gallup. [Lumina Foundation–Gallup survey, October 2025]

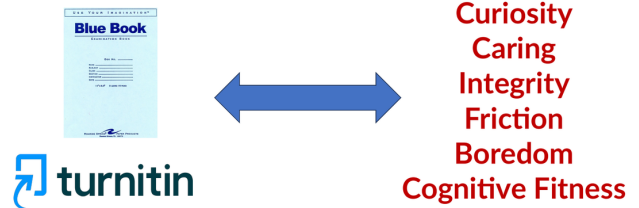
<https://news.gallup.com/poll/704087/college-students-weigh-impact-majors-careers.aspx>



Blake, S. (2025, July 31). Half of students say AI is most important skill they'll learn in college. *Newsweek*. [reports a Grammarly/Talker Research survey of 2,000 U.S. college students, July 2025] <https://www.newsweek.com/half-students-say-ai-most-important-skill-theyll-learn-college-2107250>



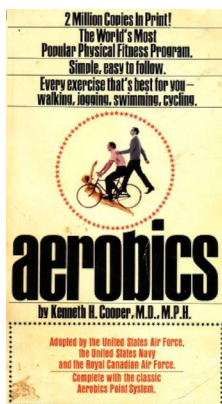
Are we solving the wrong problem?



COGNITIVE FITNESS



Kranzberg, M. (1986). Technology and history: "Kranzberg's laws." Technology and Culture, 27(3), 544–560.



Dr. Kenneth H. Cooper
(1968)

**"The Streets Are
Going to Be Full of
Dead Joggers"**

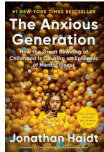
Cooper, K. H. (1968). Aerobics. M. Evans. [the "dead joggers" line is from a review from the period, not a quotation from the book]

IS COGNITIVE FITNESS IS DECLINING?

↓ **Poorer
Cognition**

📱 **Lower
Attention**

😞 **Worse
Mental
Health**



When is AI junk food for the mind?



Healthy Minds Network. (2024). Healthy Minds Study 2023–2024 data report. <https://healthymindsnetwork.org/research/data-for-researchers>

Bratsberg, B., & Røgeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. PNAS, 115(26), 6674–6678. <https://doi.org/10.1073/pnas.1718793115>

Nguyen, L., et al. (2025). Feeds, feelings, and focus: A systematic review and meta-analysis examining the cognitive and mental health correlates of short-form video use. Psychological Bulletin, 151(9), 1125–1146. [71 studies, 98,299 participants] <https://psycnet.apa.org/fulltext/2026-89350-001.html>

Mark, G. (2023). Attention span. Hanover Square Press.

Haidt, J. (2024). The anxious generation. Penguin Press. [contested — the causal claim is disputed by several researchers]

COGNITIVE DELEGATION TO AI CAN

↓ **REDUCE**
Cognitive Effort

⚠️ **DIMINISH**
Rewarding Tasks

↑ **RAISE**
Expectations

🔄 **INCREASE**
Context Switching

Lodge, J. M., & Loble, L. (2026). Artificial intelligence, cognitive offloading and implications for education. University of Technology Sydney. <https://doi.org/10.71741/4pyxmbnjq.31302475.v2>

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. Smart Learning Environments, 11(1), 28. <https://doi.org/10.1186/s40561-024-00316-7>

UNDERSTANDING COGNITIVE DELEGATION

*is essential for the
future of human cognitive fitness*



**Which tasks
should AI do?**



**When should I NOT
use AI?**

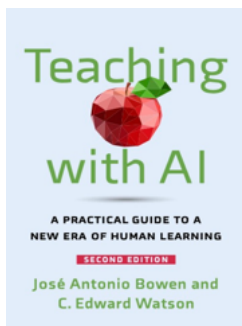


**From Professor of Content
To Cognitive Coach**

Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman. [mastery experiences build self-efficacy]

Wu, H., Ni, H., He, J., & Wee, E. H. (2026). AI anxiety and AI dependence among undergraduates: A moderated mediation model of AI self-efficacy and AI literacy. Frontiers in Psychology, 17, 1884382. <https://doi.org/10.3389/fpsyg.2026.1884382>

CITATIONS & RESOURCES



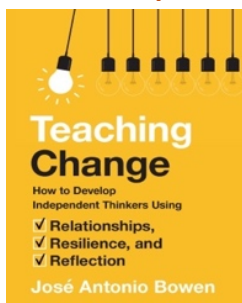
► [WeTeachWithAI.com](https://www.weteachwithai.com) contains many many more resources, prompts, assignments and links. It is updated weekly if not daily.

► **Buy the book at 30% off!**

► [Teaching with AI at Johns Hopkins University Press — code HTAI25 for 30% off](#)

► [Teaching with AI on Amazon](#)

► If you don't want to use AI, but want students to want to do things the hard way, then read **Teaching Change: How to Develop Independent Thinkers Using Relationships, Resilience and Reflection.**



► [Teaching Change at Johns Hopkins University Press — code HTWN for 30% off](#)

► [Teaching Change on Amazon \(often 25% off\)](#)