

University of Houston Law Center,
AI, Law, and the Legal Profession
October 6, 2026

Empirical Evidence on Generative AI and Lawyering

Daniel Schwarcz

Fredrikson & Byron Professor of Law
University of Minnesota Law School
schwarcz@umn.edu



Roadmap

1 The landscape

Four years of AI in law, 2022–2026.

2 Evidence on Productivity

Benchmarks and randomized controlled trials.

3 The risks

Hallucination, confidentiality, skill erosion.

4 When to use AI

A simple rule for delegating legal work.

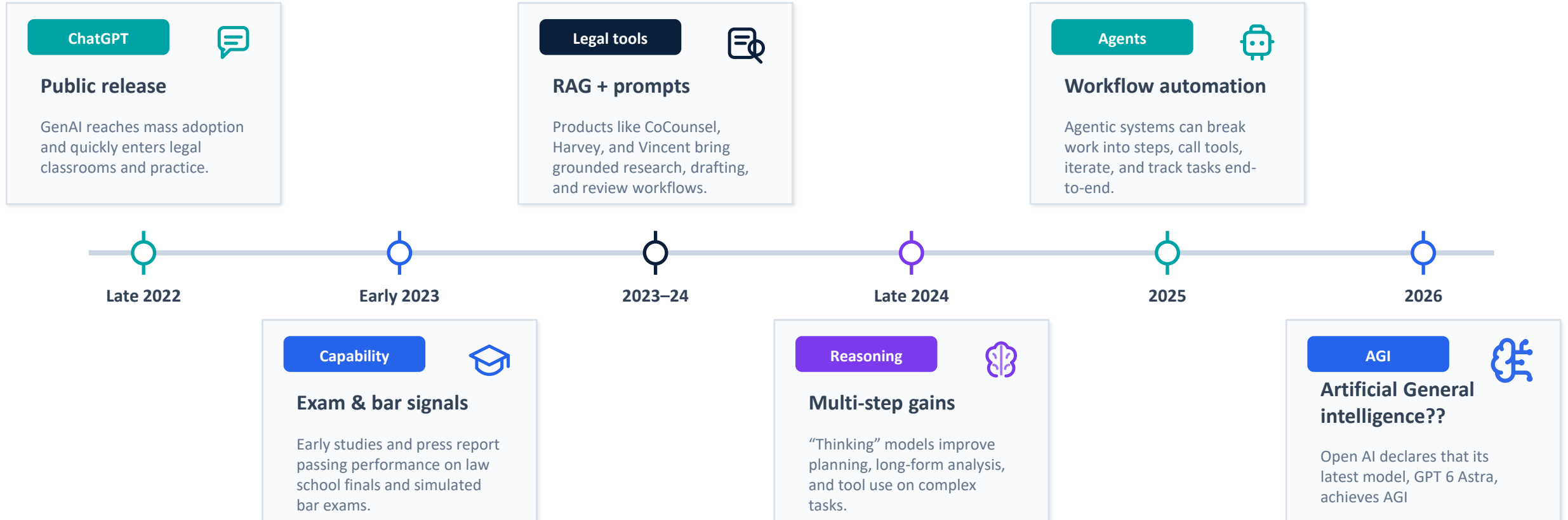
5 How to use AI

Techniques lawyers already excel at.

6 Legal education

Grading, feedback, and teaching with AI.

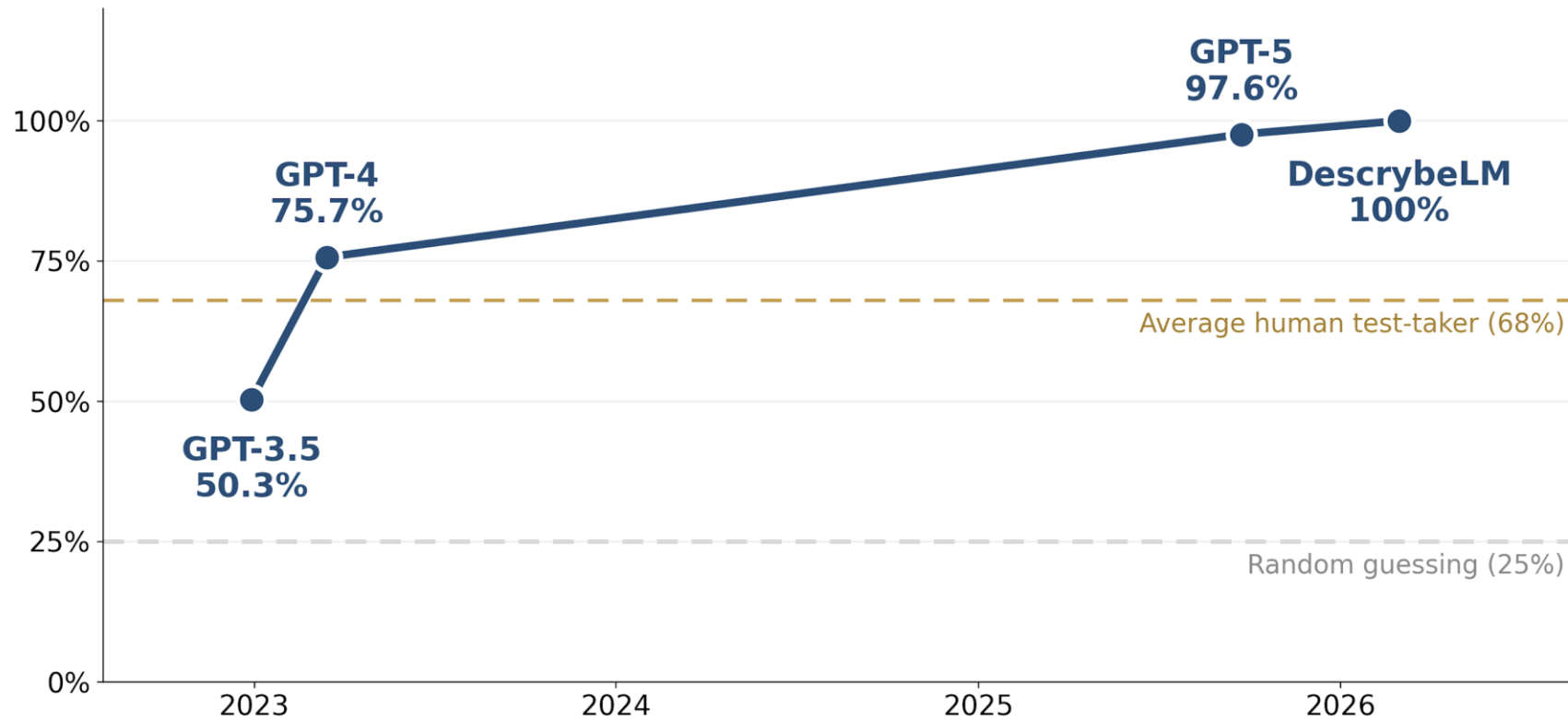
Four years of acceleration



Legal benchmarks show rapid gains—but only on bounded tasks

AI Goes from Failing to Acing the Bar Exam

Percent correct on the multiple-choice section (MBE), 2022–2026



Sources: Bommarito & Katz (2022); Katz et al. (2023); Stubenberg et al. (2025); Descrybe (2026, vendor-run benchmark).

Our method: randomized controlled trials



Participants

Upper-level UMN law students.



Tasks

Realistic junior-lawyer work — legal memos, complaint analysis, NDA drafting.



Randomization

Each group uses AI on some tasks and works unaided on others.



Grading

Blind grading by human experts, with time on task recorded.

GPT-4 saves serious time

Figure 5: Time Distributions with and Without AI – Complaint Drafting

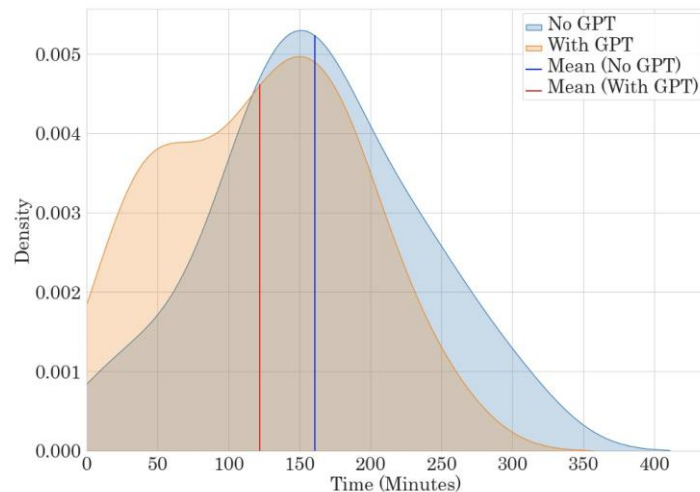


Figure 6: Time Distributions with and Without AI – Contract Drafting

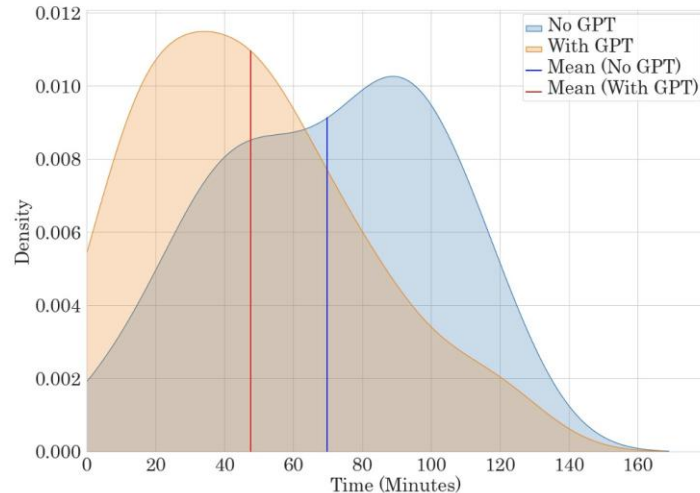


Figure 7: Time Distributions with and Without AI – Employee Handbook

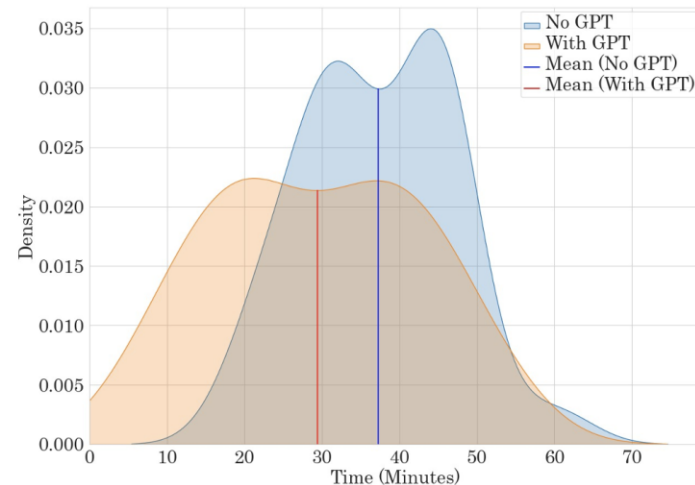


Figure 8: Time Distributions with and Without AI – Client Memo

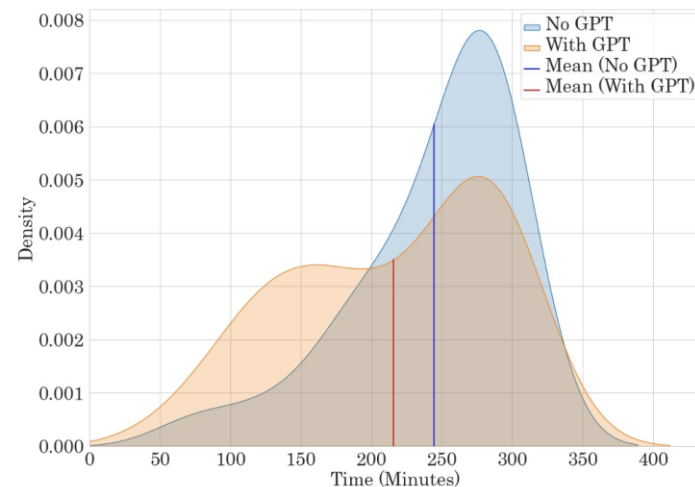


Figure 9: Task 1 vs. Task 2 Grades – Complaint Drafting

AI Reasoning models (GPT o1) and RAG (Vincent) save serious time and raise average quality of work



Figure 1: Task 1 Score Density

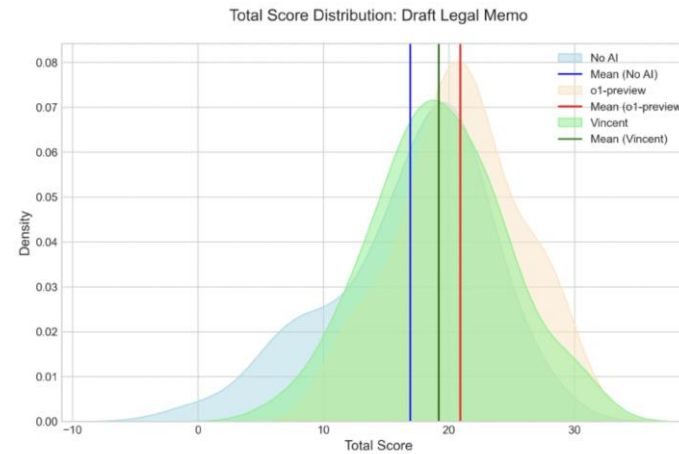


Figure 2: Task 2 Score Density

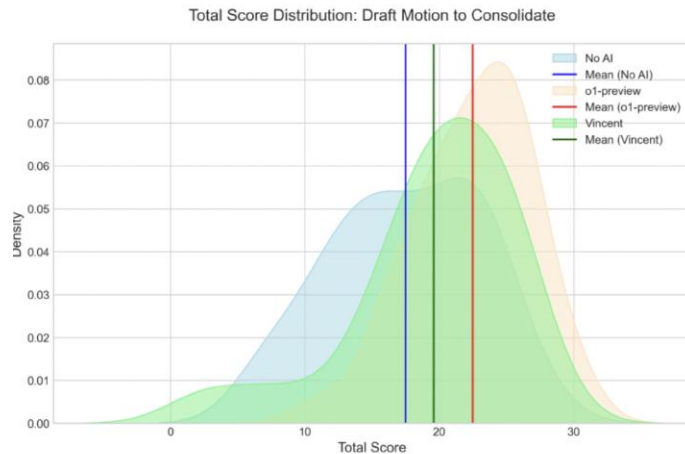


Figure 5: Task 5 Score Density

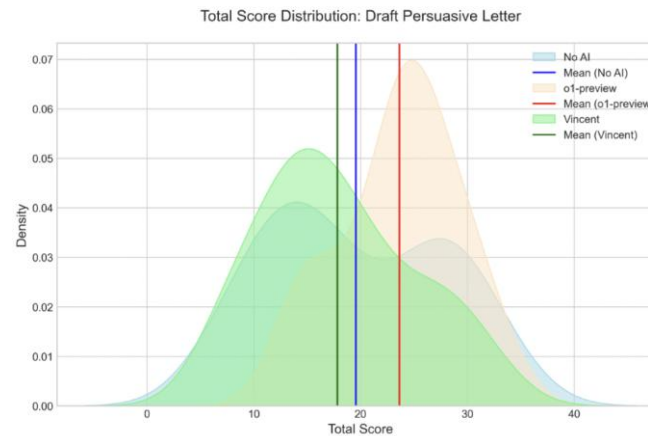


Figure 6: Task 6 Score Density

Source: Schwarcz, Manning, Barry, Cleveland, Prescott & Rich, *AI-Powered Lawyering 3 J. Law & Empirical Analysis* 220 (2026).

Three risks of AI in legal work



Hallucination

Models can invent authority—even when connected to retrieval.

Response: verify every proposition and citation.



Confidentiality

Consumer tools may retain or reuse what you submit.

Response: follow approved-tool and data-handling rules.



Crowding out reasoning

Short term: heavy reliance leaves lawyers unable to explain, extend, or refine their own work.

Long term: Juniors who outsource problem-solving may never build core lawyering skills.

Testing short-term skill-erosion risk



Participants

Upper-level law students.



Three tasks

(1) Summarize rules on servitudes → (2) recall quiz → (3) apply the rules to new facts.



Groups

Control: no AI on any task. Treatment: AI on task 1 only.



Grading

Blind grading by human experts, with time on task recorded.



Hypotheses

AI wins the assisted task (1); control wins recall (2) and reasoning (3).

AI boosts the assisted task...

TABLE 1: SUMMARY TASK, HOLISTIC PERFORMANCE

	Control Mean (Std. Dev.)	AI-Assisted Mean (Std. Dev.)	Effect (% Change)	P-Value
Substantive Quality	1.95 (0.92)	3.22 (1.06)	+1.27 (+65.12%)	<0.001***
Organization	2.29 (0.96)	3.84 (1.09)	+1.55 (+67.9%)	<0.001***
Polish	2.39 (0.94)	3.54 (1.01)	+1.15 (+48.12%)	<0.001***

Note: P-values estimated using a two-tailed t-test.

TABLE 2: SUMMARY TASK, TIME TO COMPLETION

	Control Group Mean (Std. Dev.)	AI-Assisted Group Mean (Std. Dev.)	Effect (% Change)	P-Value
Minutes	69.63 (7.27)	60.70 (12.25)	-8.92 (-12.81%)	<0.001***

Note: P-value estimated using a two-tailed t-test.

Consistent with Hypothesis 1

The AI-assisted group outperforms the control group on Task 1 — and finishes faster.

...without eroding recall or reasoning

TABLE 2: RECALL QUIZ, OVERALL PERFORMANCE

	Control Group Mean (Std. Dev.)	AI-Assisted Group Mean (Std. Dev.)	Effect (% Change)	P-Value
Recall Quiz	3.86 (1.37)	3.88 (1.21)	+0.02 (+0.05%)	0.93

Note: P-value estimated using a two-tailed t-test.

TABLE 3: REASONING TASK, HOLISTIC PERFORMANCE

+

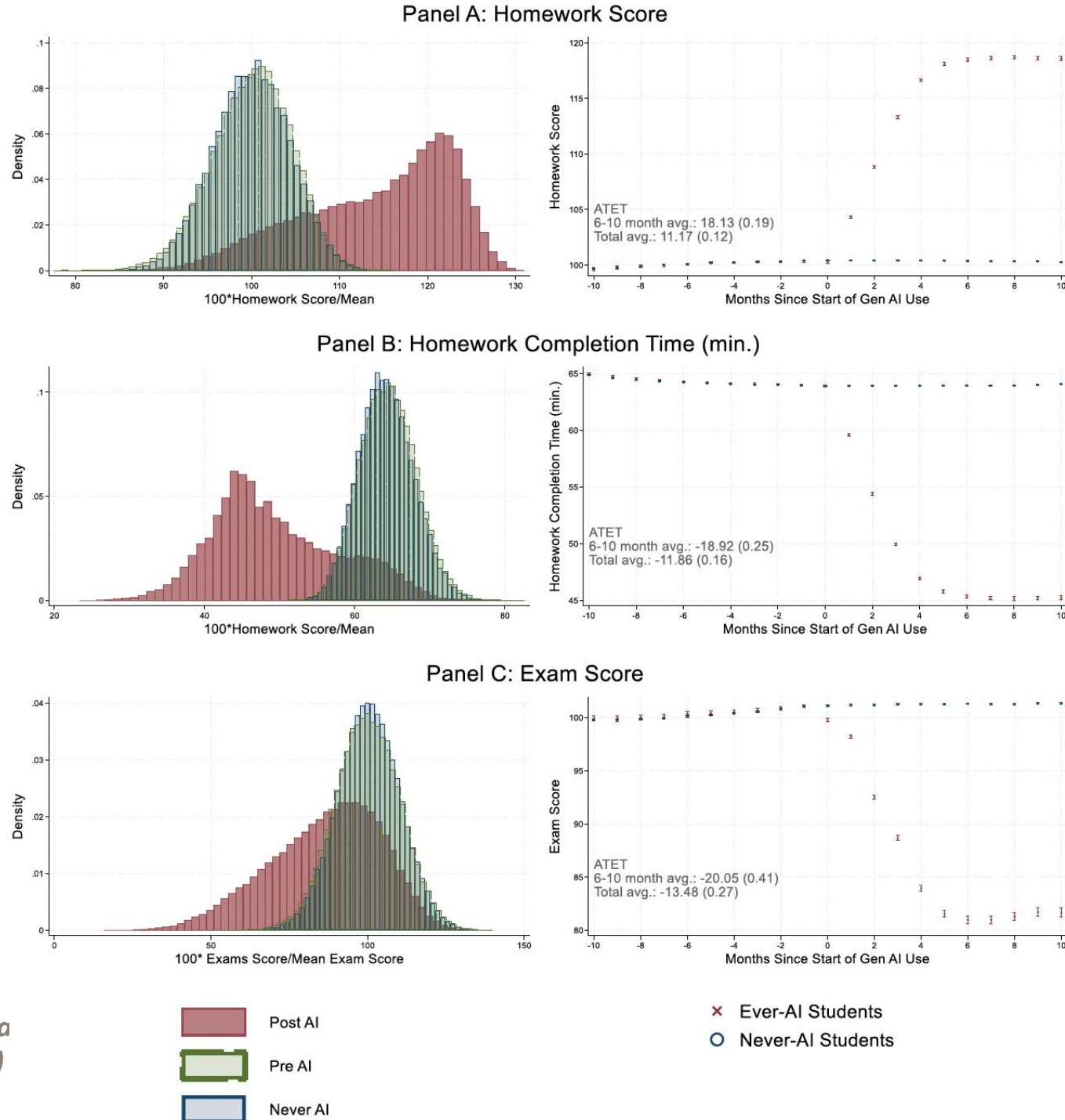
	Control Mean (Std. Dev.)	AI-Assisted Mean (Std. Dev.)	Effect (% Change)	P-Value
Focus and Emphasis	1.76 (1.02)	2.12 (1.08)	+0.36 (+20.45%)	0.10
Depth of Analysis	1.51 (0.87)	1.92 (1.03)	+0.41 (+27.15%)	0.04*
Organization	2.02 (0.99)	2.52 (0.84)	+0.50 (+24.75%)	0.01*
Polish	2.10 (0.92)	2.48 (0.84)	+0.38 (+18.09%)	0.04*

Note: P-values estimated using a two-tailed t-test.

Contrary to Hypotheses 2 & 3

The AI group matches the control group on recall — and outperforms it on reasoning.

But...Does AI use stifle development of long-term (legal) reasoning skills?



Source: Stromberg, Lei & Wu The Generative AI Learning Pena Chinese Secondary Education, unpublished manuscript (2026)

A simple rule for delegating to AI

Use AI when you can confidently **assess, adapt, explain, and build on** its output.



Reach for AI

- Improve this paragraph
- Find X in document Y
- Suggest clauses for a contract
- Flag ambiguities
- Simplify a document you understand
- Summarize a legal area you know
- Adapt a familiar template to a new context



Think for yourself first

- Draft a novel legal argument
- Answer a complex, unfamiliar research question
- Explain a case you don't understand
- Summarize a long, unfamiliar document — except as a first pass

Lawyers are built for this

Prompting well takes precision, context, and relentless follow-up — **the core skills of legal reasoning.**

CORE HABITS



Be precise

Say exactly what you want — and what you don't.



Give context

Share the background a colleague would need.



Keep probing

Follow up to surface assumptions and close gaps.

KEY TECHNIQUES



Attach the sources

Cases, statutes, guidance — ground the model in the record.



Show an example

A model of the work product you want back.



Experiment freely

Trying is cheap: compare tools, prompts, and approaches.

What this means for law schools



Early use by students is risky

Unstructured AI use by students in legal reasoning classes can undermine learning.



AI fluency is essential in long-term

Students must learn to use AI well — when to teach it is the open question.



Legal Reasoning Skills are Key to Building AI Fluency

Lawyers who can distinguish good and bad work, anticipate what they don't know, use language precisely, and spot assumptions are best positioned to leverage AI.

Questions?

Daniel Schwarcz

schwarcz@umn.edu

Learn More:

Schwarcz, Manning, Barry, Cleveland, Prescott & Rich,
AI-Powered Lawyering 3 J. Law & Empirical Analysis 220 (2026), <https://ssrn.com/abstract=5162111>

Bednar, Cleveland, Erbsen & Schwarcz,
Artificial Intelligence and Human Legal Reasoning, Northwestern Law Review (forthcoming, 2027),
<https://ssrn.com/abstract=6525800>

Choi, Monahan & Schwarcz,
Lawyering in the Age of Artificial Intelligence, 109 Minn. L. Rev. 147 (2024), <https://ssrn.com/abstract=4626276>

Cope, Frankenreiter, Hirst, Posner, Schwarcz, & Thorley
Grading Machines: Can AI Exam-Grading Replace Law Professors?, 3 J. Law & Empirical Analysis 2 (2026),
<https://ssrn.com/abstract=5851362>