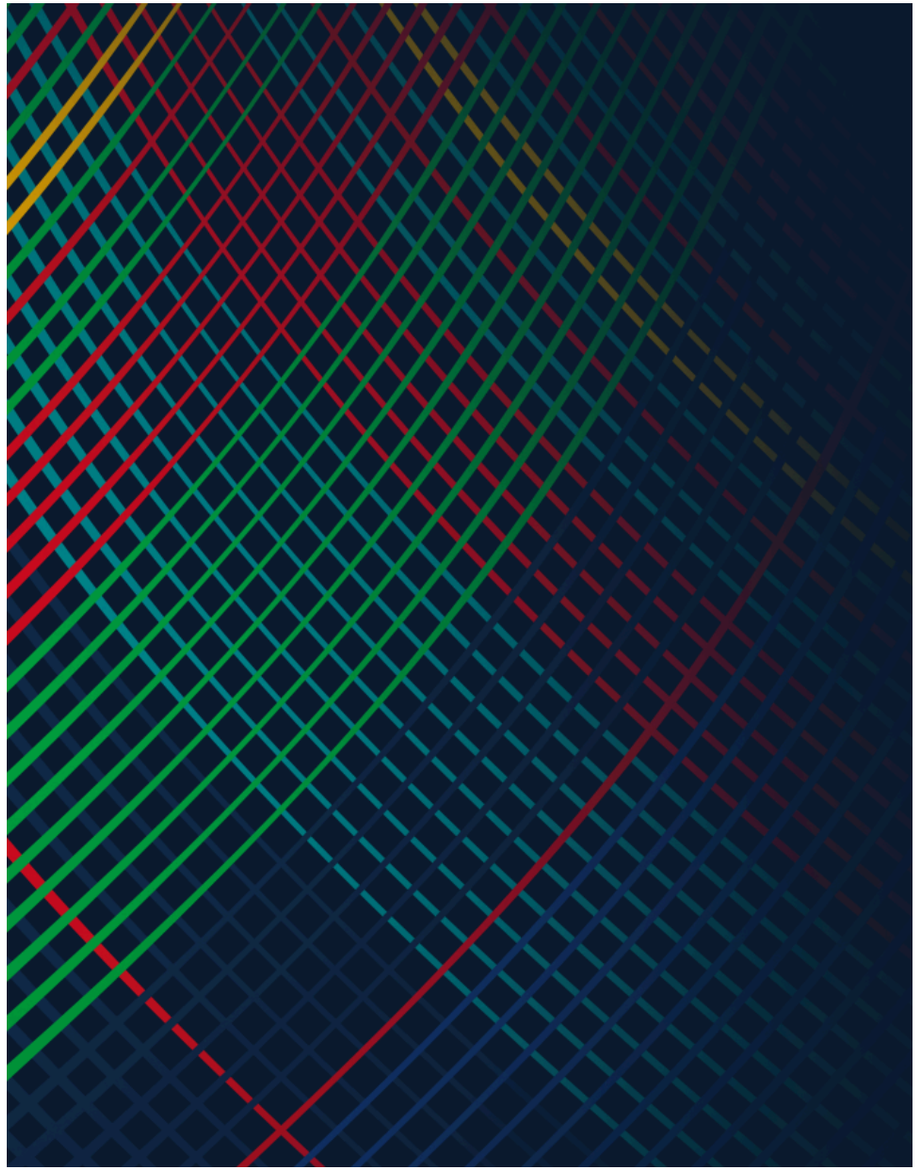


Writing to learn or learning to write?

Writing in the statistics curriculum and the role of LLMs

Alex Reinhart
areinhar@stat.cmu.edu
TeachStat Group, Statistics & Data Science, Carnegie Mellon University

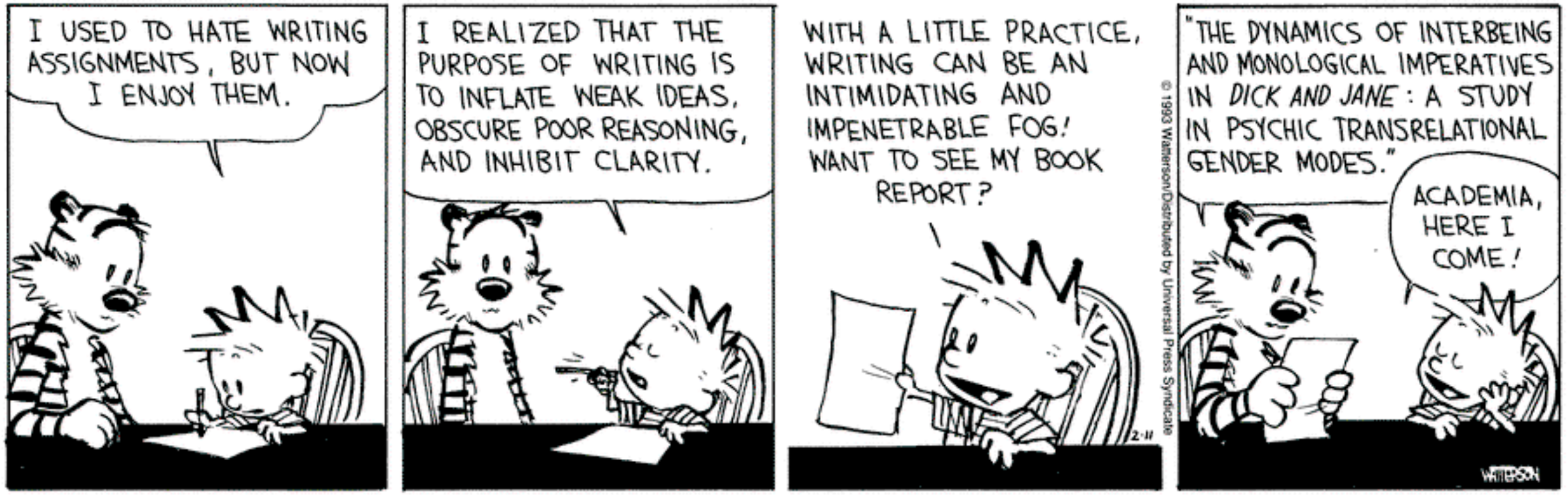
May 20, 2025

Main takeaways

- What is the role of writing in the curriculum now that we have large language models?
- I have no idea
- But we can study how students, experts, and LLMs write to better understand what happens when students use LLMs to write like experts
- It turns out that LLMs don’t knit the same sweaters experts do

What is writing for?

The purpose of writing



Writing to learn

A writing assignment can help students learn by:

- Forcing them to organize their ideas
- Making misconceptions visible to the instructor
- Providing opportunities for feedback

Assignments are often “inauthentic”, but helpful nonetheless

Examples:

- Concept maps
- Entry and exit tickets
- Minute papers
- Learning journals

(Woodard, Lee, and Woodard 2020)

☞ If students outsource the tasks to LLMs, do they get the benefits?

Writing in the disciplines

Writing can also be a core professional competency:

- Publishing academic papers
- Submitting reports to management
- Providing deliverables to consulting clients

These tasks require *authentic* practice and feedback

Your Goals

You have been hired by an eccentric billionaire who is willing to do whatever it takes to become immortal. The billionaire, Preston Jorgensen, would like you to review the AnAge data to determine what factors appear to be related to lifespan, so he can fund research projects that try to use those factors to attain immortality.

Specifically, Jorgensen wants you to answer three questions:

(i) Some research has suggested that slowing the metabolic rate—for example, by intermittent fasting—may increase lifespan. Is this supported by the evidence? Develop a model for lifespan using metabolic rate, and interpret what it means.

(ii) Is the relationship between metabolic rate and lifespan nonlinear, even after transformations? Use a nonparametric model to determine if a nonlinear fit is more appropriate.

(iii) Jorgensen would like a pet to keep him company during his immortality. He thinks the crab-eating raccoon (*Procyon cancrivorus*, HAGRID = 1898) would be a good companion. If he pays for research to reduce its metabolic rate by 50%, what does your model estimate for the mean lifespan of an animal with the crab-eating raccoon's characteristics but a 50% smaller metabolic rate?

☞ If students outsource the tasks to LLMs, will they learn to write like professionals?

How students develop writing skills

There are many kinds of writing

- “Writing” isn’t one task; students learn different kinds:
  - Business memos
  - Lab reports
  - Data analysis reports
  - Argumentative essays
  - Client presentations
- Each has unique features that students must learn
- These features are also essential for professional writing
- Students increasingly use large language models to help

Writing skill development takes time

Students tend to start with an involved, interactional style and move gradually to a more informational style:

Involved style:

- More pronouns (*I*, *we*, *everyone*)
- Private verbs (*think*)
- Contractions (*can't*, *that's*)
- More disjointed
- Focuses on personal stance

Informational style:

- Lots of nouns and noun structures (*a parametrized covariance structure that needs to be estimated*)
- Frequent nominalizations (nouns formed from verbs and other parts of speech)
- More informationally dense
- More "compressed"

(Staples et al., 2016; Gardner et al., 2018; Carter, 1990)

Human-AI Parallel English corpus

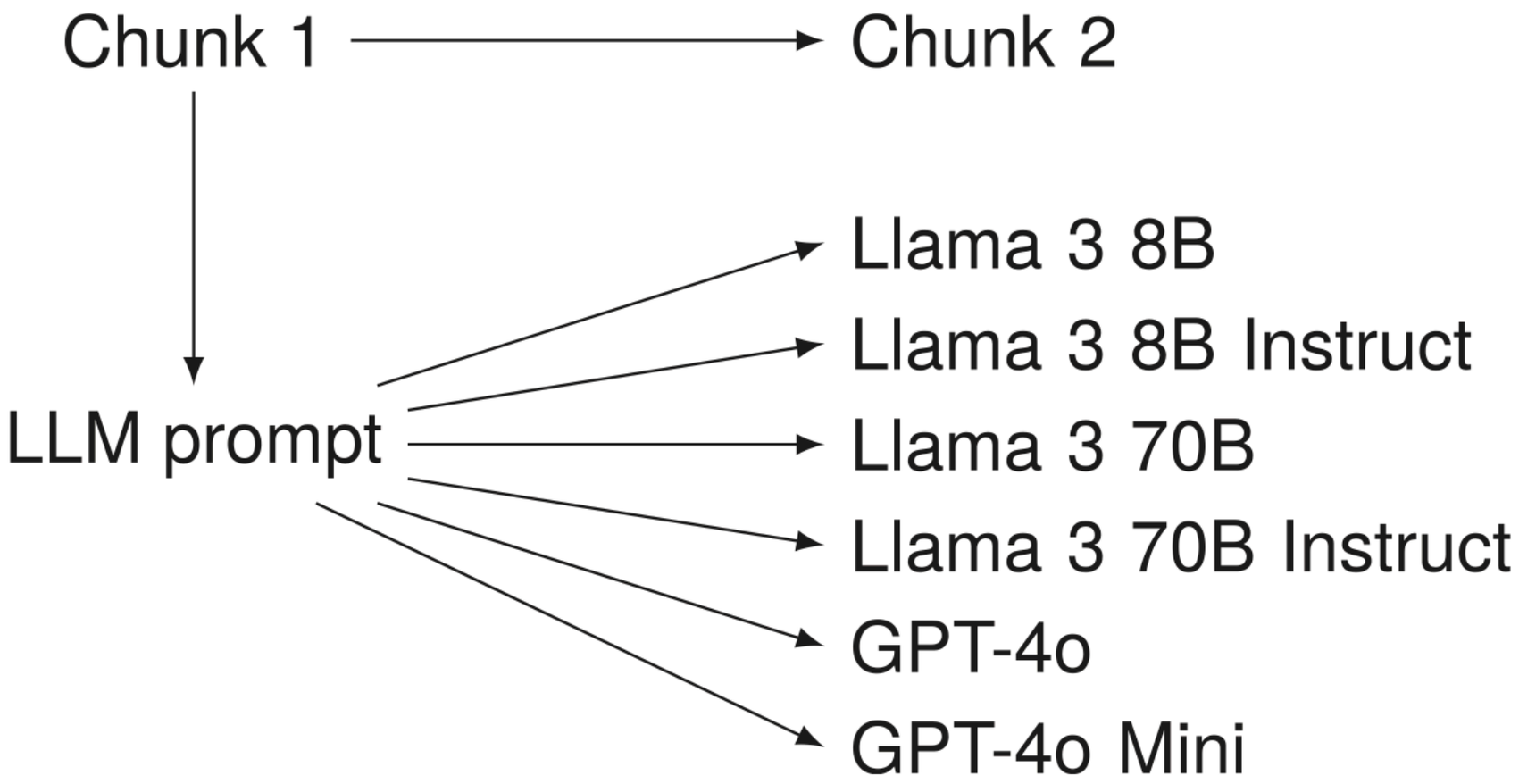
The Human-AI Parallel English Corpus

We constructed the HAP-E corpus of 12,000 human texts:

- Open-access academic papers
- News articles
- Public-domain novels and short stories
- Podcast transcriptions
- Blog posts
- TV and movie scripts

We used these to generate LLM output from OpenAI's GPT-4o and Meta's Llama 3

Parallel construction allows direct style comparison



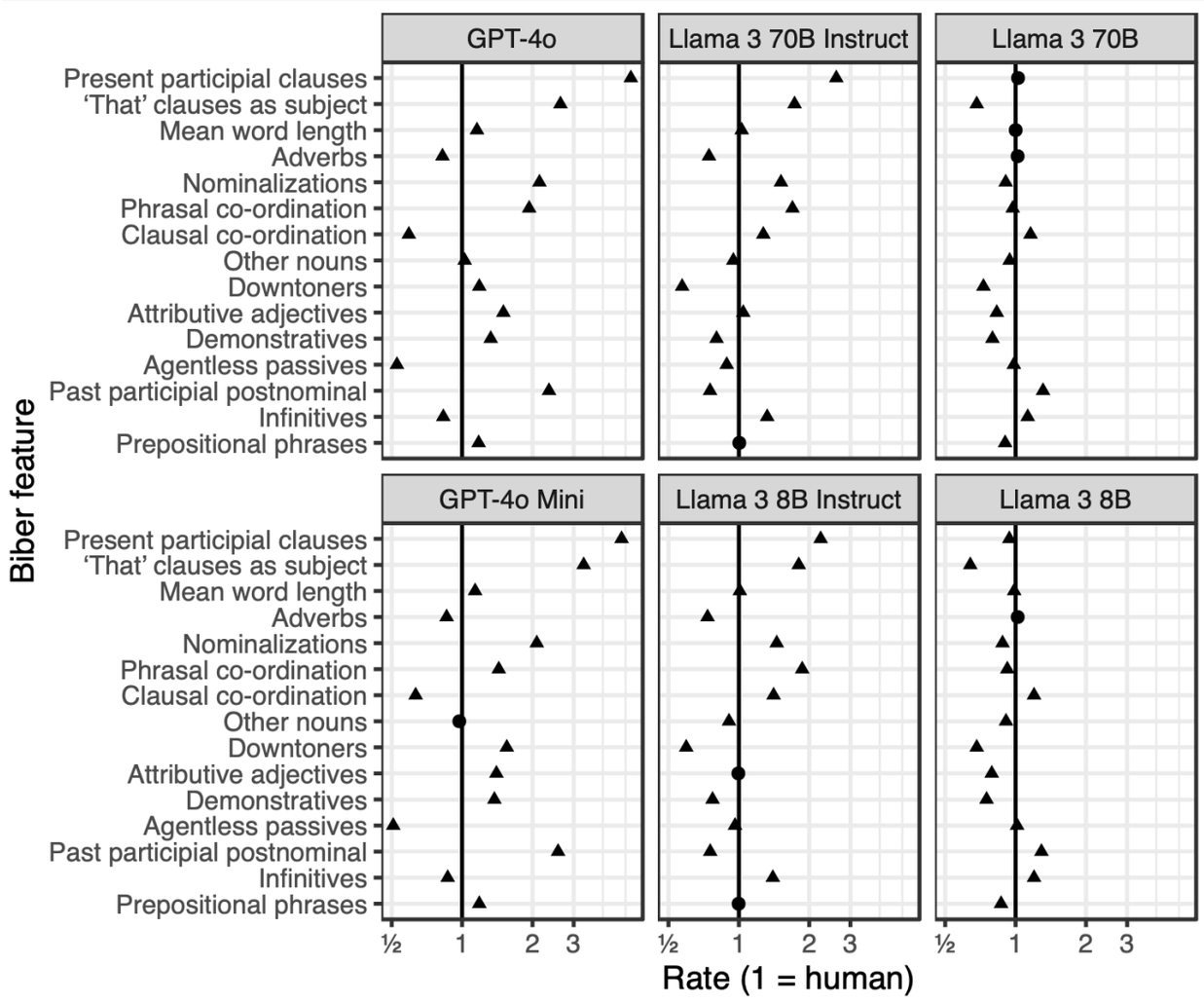
Biber's features reflect style and genre

Douglas Biber constructed a set of 67 linguistic features to explore register variation:

- Verb tenses
- Parts of speech
- Nominal forms (e.g., nominalizations, like "justification" or "abandonment")
- Past and present participles
- Many types of sentence structural features
- Hedges, amplifiers, emphatics, downtoners
- Contractions

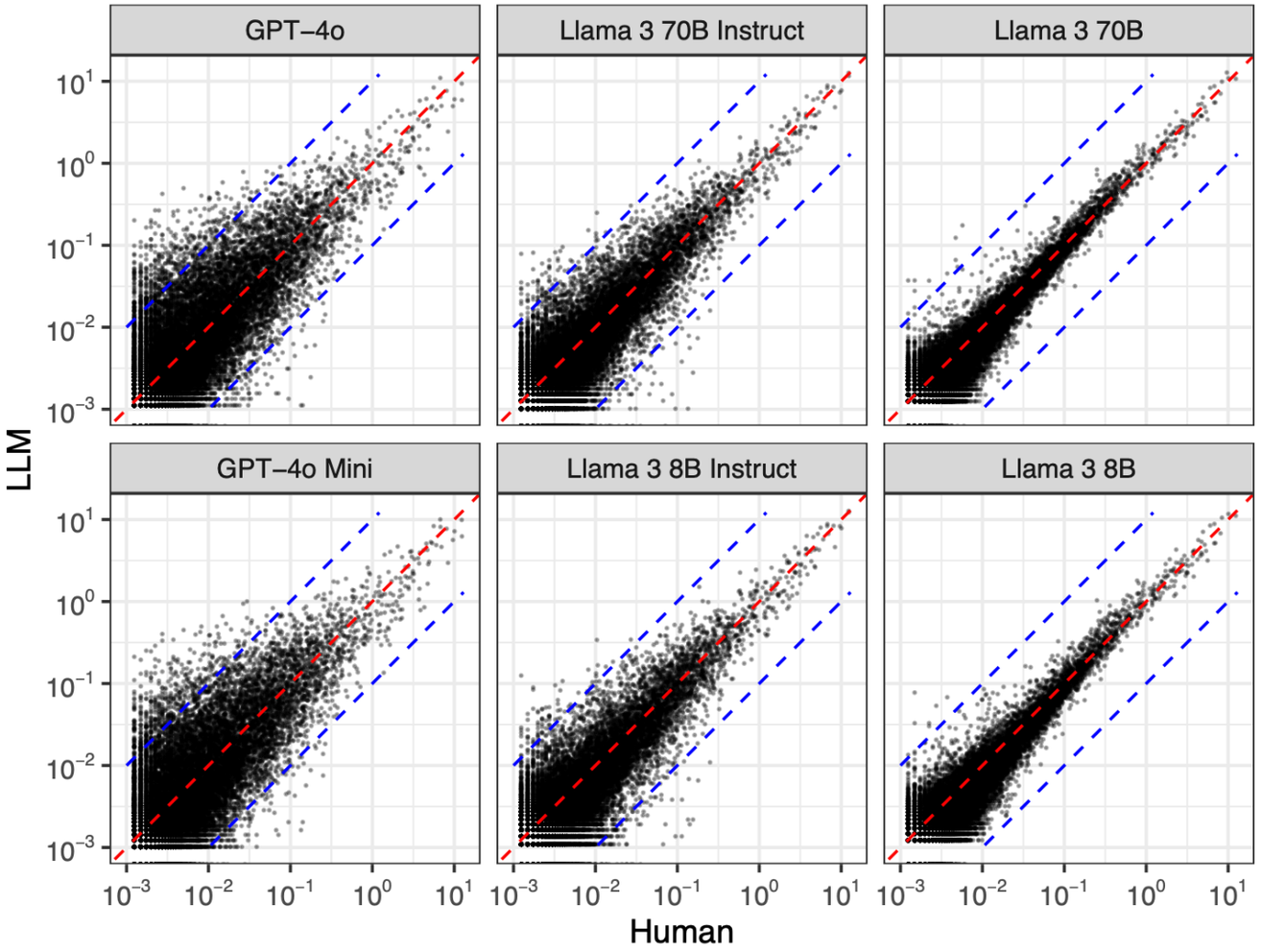
These have been widely used to study variation in human writing between registers, from casual speech to formal academic writing

LLM Biber features strongly differ from humans



- Differences larger for instruction-tuned LLMs
- Features favor information density and greater formality
- Appear to be different between AI companies

LLM vocabulary use strongly differs from humans



GPT-4o's favorite words:

- camaraderie (162×)
- tapestry (155×)
- intricate (119×)
- underscore (107×)

Llama 3's:

- unease (63×)
- palpable (47×)
- continuation (29×)
- shoutout (28×)

What about statistical writing?

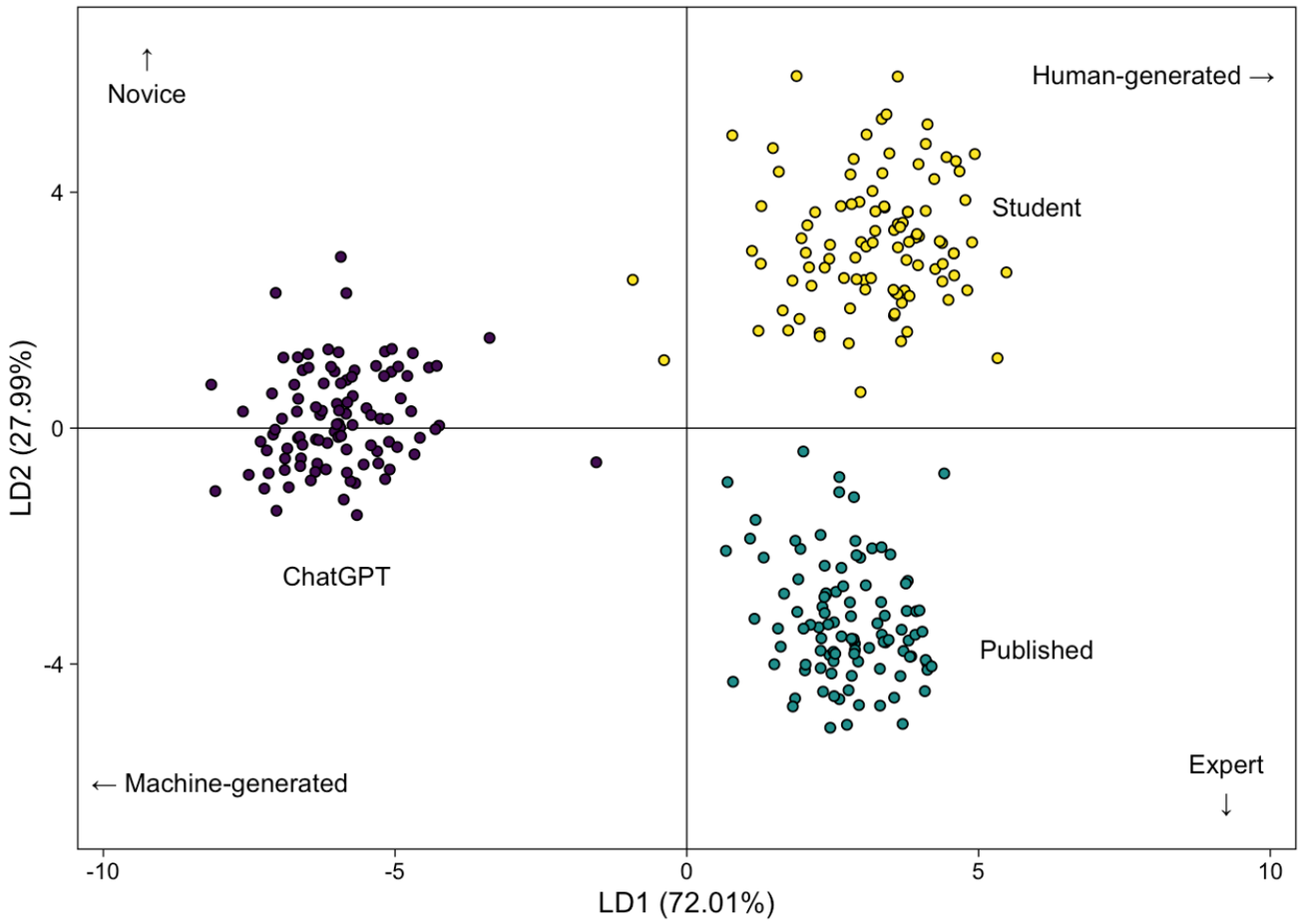
How well does ChatGPT write statistical reports?

We (DeLuca et al., 2025) asked ChatGPT to complete two tasks:

- Write introductions for intro-level data analysis report (GPT-3.5)
- Write a complete data analysis report, *including writing and running Python code*, for a PhD-level regression project (GPT-4)

For comparison, we compiled a corpus of introductions from published scientific writing (Elsevier open access) and from student writing

ChatGPT's introductions are stylistically distinct



- Novices are more emphatic, focus more on people and concrete actions, use fewer distinct verbs
- Experts are more abstract, focus more on information than on agents
- ChatGPT uses more nouns, longer words, more information density, less modulation of confidence

ChatGPT's data analysis is... adequate

- ChatGPT is startling good at writing and debugging Python code for data analysis
- Its complete data analysis report was well-written but analytically uninspired
- ...and some of its results were hallucinated
- It even replicates human bad habits, like noting a $p = 0.053$ result as "marginally significant"

Conclusion

LLMs may bypass the practice that forms skills

- In writing-to-learn, the act of writing helps students learn
  - If you use an LLM to write your minute paper or concept map, do you benefit from it?
  - We may need entirely new activities to achieve the same goals
- For professional skills, LLM writing is not a direct substitute for expert writing
  - How can students learn to write more like experts?
  - There are opportunities to use LLMs for feedback and growth, but not if students outsource all work to them

☞ We must reassess the skills students need and devise new ways to teach them

Thank you

Thanks to our entire research group and the coauthors below:

- Reinhart, A., Markey, B., Laudenbach, M., Pantusen, K., Yurko, Y., Weinberg, G & Brown, D. W. (2025). Do LLMs write like humans? Variation in grammatical and rhetorical styles. *Proceedings of the National Academy of Sciences*, 122(8), e2422455122.
- DeLuca, L., Reinhart, A., Weinberg, G., Laudenbach, M., Miller, S., & Brown, D. W. (2025). Developing students' statistical expertise through writing in the age of AI. *Journal of Statistics and Data Science Education*, in press.

Contact me: areinhar@stat.cmu.edu