

# Slaying Educational Zombie Claims with Empiricism and AI: Is the Pen Truly Mightier than the Laptop?

David J. Therriault<sup>1</sup>, Carina A. Swenson<sup>1</sup>, Nitya S. Kodali<sup>1</sup>, and Lise Abrams<sup>2</sup>

<sup>1</sup>School of Human Development and Organizational Studies in Education, University of Florida

<sup>2</sup>Pomona College

## Abstract

Zombie claims in education are characterized as persistent myths or misnomers that, once created, continue to influence the learning environment, often without empirical grounding (De Bruyckere et al., 2015; Sinatra & Jacobson, 2019). One such claim that quickly obtained zombie status is the belief that handwritten notes are superior to laptop notes (Mueller & Oppenheimer, 2014). More recent research has produced mixed results or failed to replicate the original study (Morehead et al., 2019; Urry et al., 2021). The current study explored the differences in learning performance when participants were instructed to type or handwrite their notes in a mock class setting (i.e., after viewing a lecture). Their memory was tested a week later. Notetaking performance was compared to a control condition (sit and listen without taking notes), and participants' handwriting and typing speed/accuracy were measured as possible covariates. Notes were analyzed using NaviGator Chat to attempt to transcribe handwritten notes into electronic form. Finally, we compared the quality of student notes to a generated AI gold standard (created from the lecture transcription) and used their “notes” grade to predict exam performance.

## Background

### Longhand-Superiority effect:

- Mueller and Oppenheimer (2014):
  - Participants watched video lectures while taking notes by longhand on paper or by typing on a laptop.
  - Exam performance on conceptual questions was greater for those who took notes longhand than by laptop.
- Flanigan et al. (2024):
  - Meta-analysis shows benefit of longhand notes over laptop typing

### Academic and Media response to Mueller and Oppenheimer:

- Highly cited (Google Scholar [n > 2000])
- Media impact: NPR and The Atlantic published provocative articles on the benefits of handwriting notes.
- BUT: the longhand-superiority effect does not always replicate (e.g., Urry et al., 2021; Richardson & Lacroix, 2023), and laptop notetaking may actually be beneficial in some circumstances (e.g., Bui et al., 2013; Morehead et al., 2019)
- Few studies employ control conditions, test memory with realistic classroom delay (e.g., one week or more), or examine the notes.**

## Research Questions

- RQ#1:** Is there a difference in exam performance when notes are hand-written versus typed in a mock classroom setting when memory is tested a week later?
- RQ#2:** Does the speed of notetaking influence exam performance?
- RQ#3:** Can AI be used to transcribe handwritten notes accurately?
- RQ#4:** Can AI-graded notes be used to approximate exam performance?

## Experiment

### Participants

One hundred and fifty-seven undergraduates (N = 157; 84% female;  $M_{age}$  = 19.71;  $M_{GPA}$  = 3.51) participated for course credit. Each student was randomly assigned to one of the three conditions:

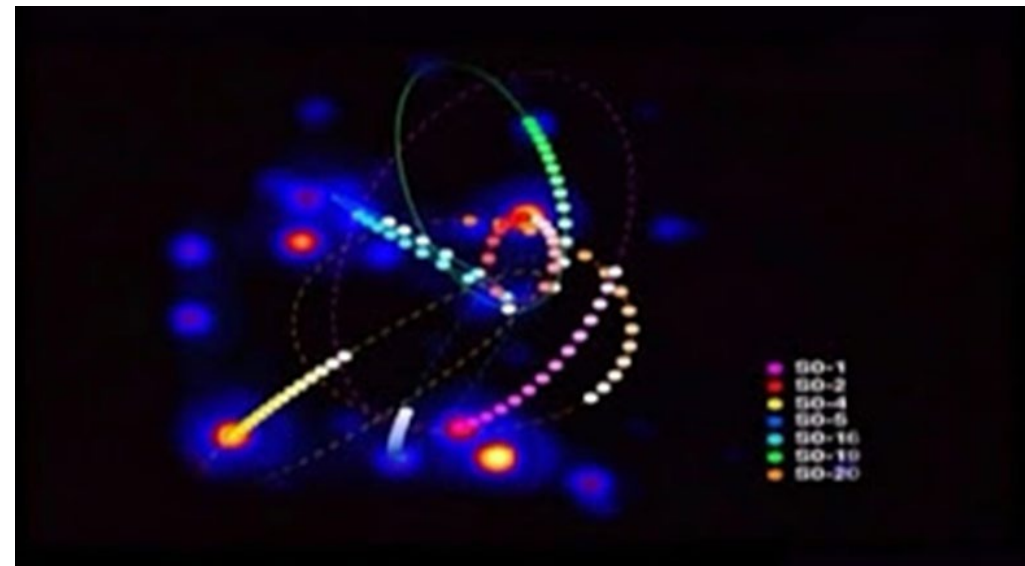
- Hand-writing notes ( $N=51$ )
- Typing notes ( $N=52$ )
- Watching lecture only (control) ( $N=54$ )

### Procedures

#### Session 1:

Participants took a pretest over the material (25 MC questions), watched a TED talk on black holes, and either took handwritten notes, typed their notes, or only watched.

**Materials:** Ted Talk, “The hunt for a supermassive black hole”, by Andrea Ghez (16 minutes)



#### Session 2:

Participants returned one week later: They were allowed to review their notes for 5 minutes, then took the same exam as the pretest. Following the exam, participants took writing and typing speed tests.

### Sample Exam Questions:

What is the Schwarzschild radius of the object.

- An aggregation of small particles orbiting a massive planet or black hole.
- The radius of binary stars calculated using the total gravitational pull.
- A virtual radius proportional to the mass of a black hole.
- The distance between the outer bounds of neighboring black holes.

In order to prove that a black hole exists, astronomers only really need to find:

- the mass of the object.
- the mass and volume of the object.
- the volume of the object.
- the volume and Schwarzschild radius of the object.

## Analyses

RQ#1: One-way ANOVA on post-test exam performance with condition as a between-subjects variable.

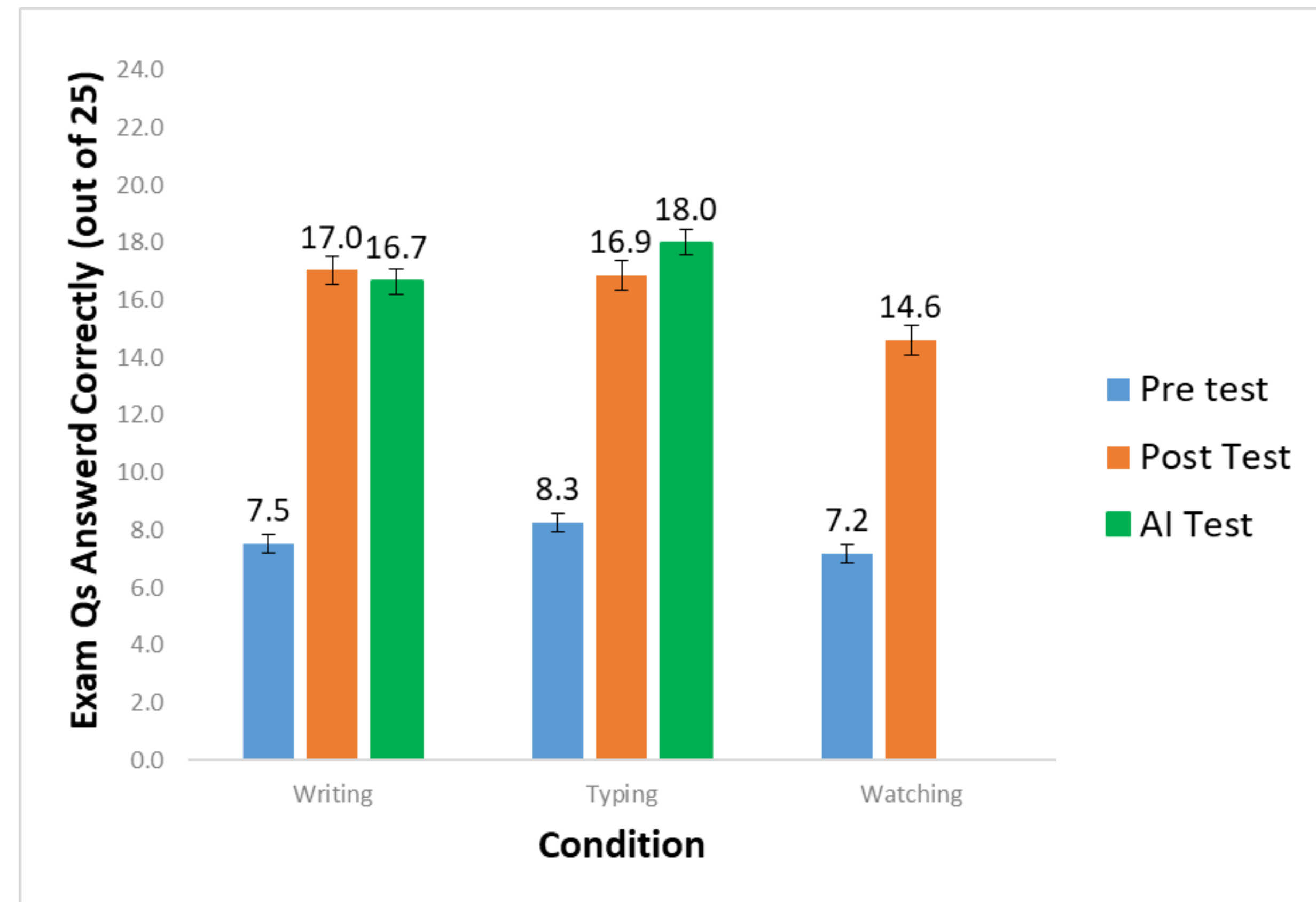
RQ#2: One-way ANCOVA on post-test exam performance with condition as a between-subjects variable and writing and typing speed as covariates.

RQ#3: Observational comparisons NaviGator

RQ#4: Follow-up within and between T-tests comparing post-test exam performance with AI-predicted scores.

## Results

**RQ#1: No.** While the ANOVA revealed a significant effect of condition ( $p = .001^*$ ), [i.e., performance in the control condition was worse than taking any kind of notes] Tukey’s HSD post hoc analyses revealed no difference between the handwriting and typing conditions,  $p = .964^{NS}$ .



Note: Error bars depict + 1 SE.

**RQ#2: No.** The speed of notetaking did not influence exam performance. Neither typing speed nor writing speed predicted exam performance,  $p \geq .116^{NS}$ .

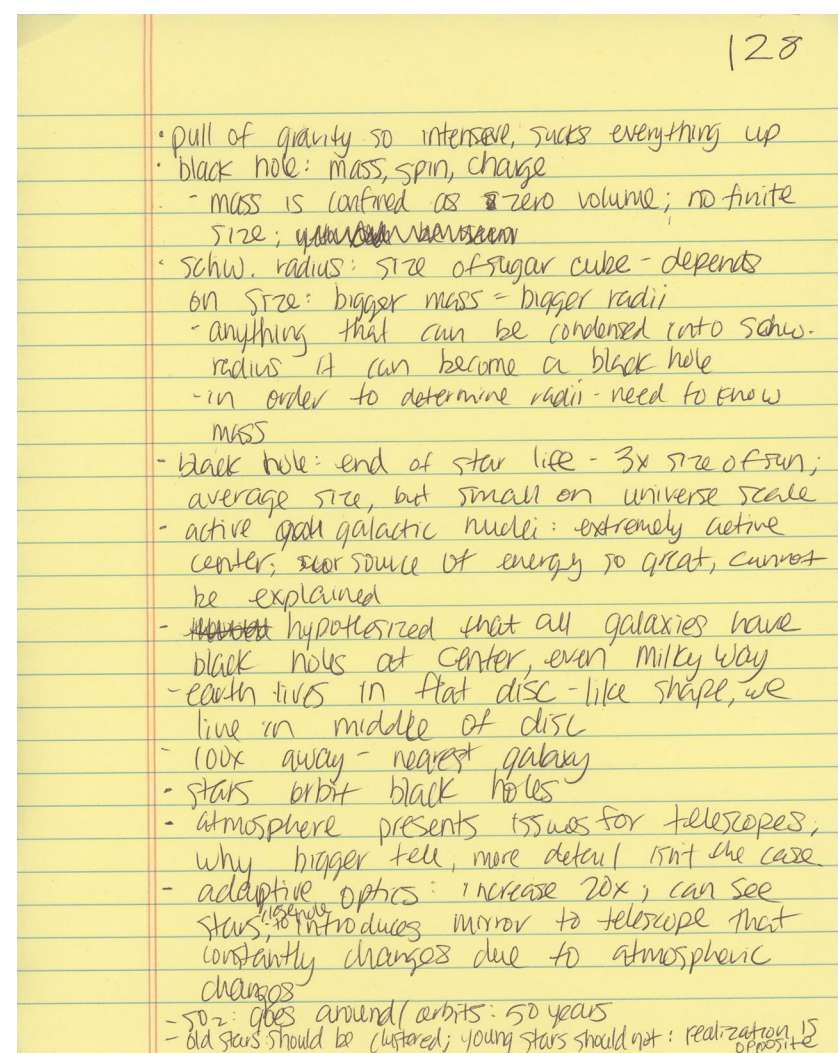
- The effect of condition remained significant,  $p = .001$ , after controlling for typing speed and writing speed.

## AI-Enabled Results

**RQ#3: Yes.** AI accurately transcribed handwritten notes with minor exceptions (adding headers, formatting, bolding, & font size errors)

### NaviGator OCR Prompt & Example Output

Please transcribe these written notes for me {image file 128}. Please use markdown formatting. Do not apply line wrapping to any paragraphs. Try to capture headings, subheadings, and any formatting, such as bold or italic but do not add any words or headings that were not originally in the notes. Omit any text that has been scribbled out. Try your best to understand the writing and produce a first draft. If anything is unclear, follow up with a question at the end.



- 128
- pull of gravity so intense, sucks everything up
  - black hole: mass, spin, charge
    - mass is confined as zero volume; no finite size
  - Schew: radius: size of sugar cube - depends on size; bigger mass = bigger radii
    - anything that can be condensed into Schew: in order to determine radii, need to know mass
  - black hole: end of star life - 3x size of sun; average size, but small on universe scale
  - active galactic nuclei: extremely active center; source of energy so great, cannot be explained
  - hypothesized that all galaxies have black holes at center, even Milky Way
  - Earth lives in flat disc-like shape; we live in middle of disc
  - 100% empty - nearest galaxy
  - stars orbit black holes
  - atmosphere presents issues for telescopes
    - why bigger telescopes more detail isn't the case
      - adaptive optics: increase 20x; can see stars, introduces mirror to telescope that constantly changes due to atmospheric changes
  - SO2: goes around orbits: 50 years
  - old stars should be clustered; young stars should not: realization is opposite

## AI-Enabled Results (Continued)

**RQ#4: Yes.** AI-graded notes can approximate exam performance (but with substantial caveats); see predicted test scores as green bars in the histogram to the left (the t-tests revealed that they did not differ from obtained scores,  $p = .172^{NS}$ ). However, we should be cautious about arguing based on a null result.

### NaviGator Grading Prediction Prompt Used

**Persona:** You are a University Professor with many years of experience teaching, grading, and researching in the area of education and psychology. **Context:** You are interested in creating a tool that grades the notes that students take when attending a lecture and how those notes might predict their performance on a test of the lecture material. **Exemplars:** Included in the attachments here are a transcript of the lecture on black holes, a gold standard set of notes to use as an exemplar of good note-taking, a copy of a student's notes, and finally, a copy of the exam the student would take. To help you understand the elements of good notes are:

- Organization: Selects, organizes and elaborates ideas expressed by speakers so that the notetaker can learn in generative and constructive ways.
- Accuracy: Information in notes are accurate to the information in the lecture
- Comprehensiveness: The notes cover all key topics (ex. definitions, examples, critical information) presented in the lecture
- Clarity: Notes include well-formed sentences and abbreviations that can be easily understood
- Critical thinking: Captures and preserves information in a form that stimulates recall of main and supporting details for the individual notetaker, so that an outside viewer may not be able to extract the same information as the notetaker because the outside viewer was not present when the notes were encoded and/or was not privy to the notetaker's thought processes and reasoning when encoding the information
- Summarization: The notes are not verbatim, but the main points and essential details are condensed

**Task and Format:** Your task is to assign the student notes a grade as a percentage (from 0%-100%) based on the quality of their notes using the elements of good note-taking in this prompt and the gold standard document provided as an attachment. Furthermore, using your assessment/grade of the student notes and the attached exam, predict how well the students would do on the exam in terms of how many questions they would get correct out of the possible 25. Provide a rationale for your grading of the notes and for the exam score prediction. If anything is unclear or additional information would be helpful, please follow up with any requests.

## Conclusions

Longhand notetaking was **not** superior to laptop notetaking (**zombie dispatched**)

- Even in more ecologically valid conditions where memory is typically assessed (in a mock class setting and after a significant testing delay).
- Consistent with mixed results previously reported within this literature.

Notetaking (of any type) increases exam performance compared to simply watching a lecture and can't be attributed to individual differences in writing or typing speed.

AI is a useful tool and could be further employed in testing the benefits of notetaking and creating apps for students to predict their learning from notetaking.

## References

- Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology*, 105(2), 299.
- Flanigan, A. E., Wheeler, J., Colliot, T., Lu, J., & Kiewra, K. A. (2024). Typed versus handwritten lecture notes and college student achievement: A meta-analysis. *Educational Psychology Review*, 36(3), 78.
- Morehead, K., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer (2014). *Educational Psychology Review*, 31, 753-780.
- Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological science*, 25(6), 1159-1168.
- Richardson, L., & Lacroix, G. (2023). Which modality results in superior recall for students: Handwriting, typing, or drawing?. *Journal of Writing Research*, 15(3), 519-540.
- Urry, H. L., Crittle, C. S., Floerke, V. A., Leonard, M. Z., Perry III, C. S., Akdilek, N., ... & Zarrow, J. E. (2021). Don't ditch the laptop just yet: A direct replication of Mueller and Oppenheimer's (2014) study 1 plus mini meta-analyses across similar studies. *Psychological Science*, 32(3), 326-339.