



# AI-Supported Question Generation for Scientific Reading

*Design and Evaluation of a Learning-Oriented Chatbot*

---

**Piumal Mahawasala** · M.Sc. Educational Technology



# About Me & Research Interests

## Interdisciplinary Background

Computer Science → Education  
→ Educational Technology

**MSc EdTech** (Germany), **M.Ed.**  
(Leadership), **BSc** (CS)

## Educator & Leader

Secondary School Teacher

School Principal (digital  
transformation initiative)

## Builder & Entrepreneur

Founder of ScienceDojo

Designed AI-supported learning  
tools & courses

## Research Interest

Human-AI interaction in learning

Scalable educational systems

MOTIVATION

# The Reading Problem in Science Education



- **Complex Texts**  
Scientific literature demands high cognitive load, domain vocabulary, and inferential reasoning.
- **Passive Reading**  
Learners frequently skim without monitoring comprehension or forming deeper connections.
- **Active Engagement Gap**  
Genuine understanding requires self-questioning — yet few learners practice it spontaneously.

# Where Current Tools Fall Short

## No Self-Questioning Habit

Most learners do not generate their own questions while reading, missing a critical metacognitive strategy for consolidating knowledge.

## Limited Cognitive Tools

Existing reading aids focus on summarization or highlighting – not on prompting the reader to think critically about content.

## Result → Shallow Learning

Passive reading leads to shallow encoding, poor retention, and an inability to transfer knowledge to new contexts.

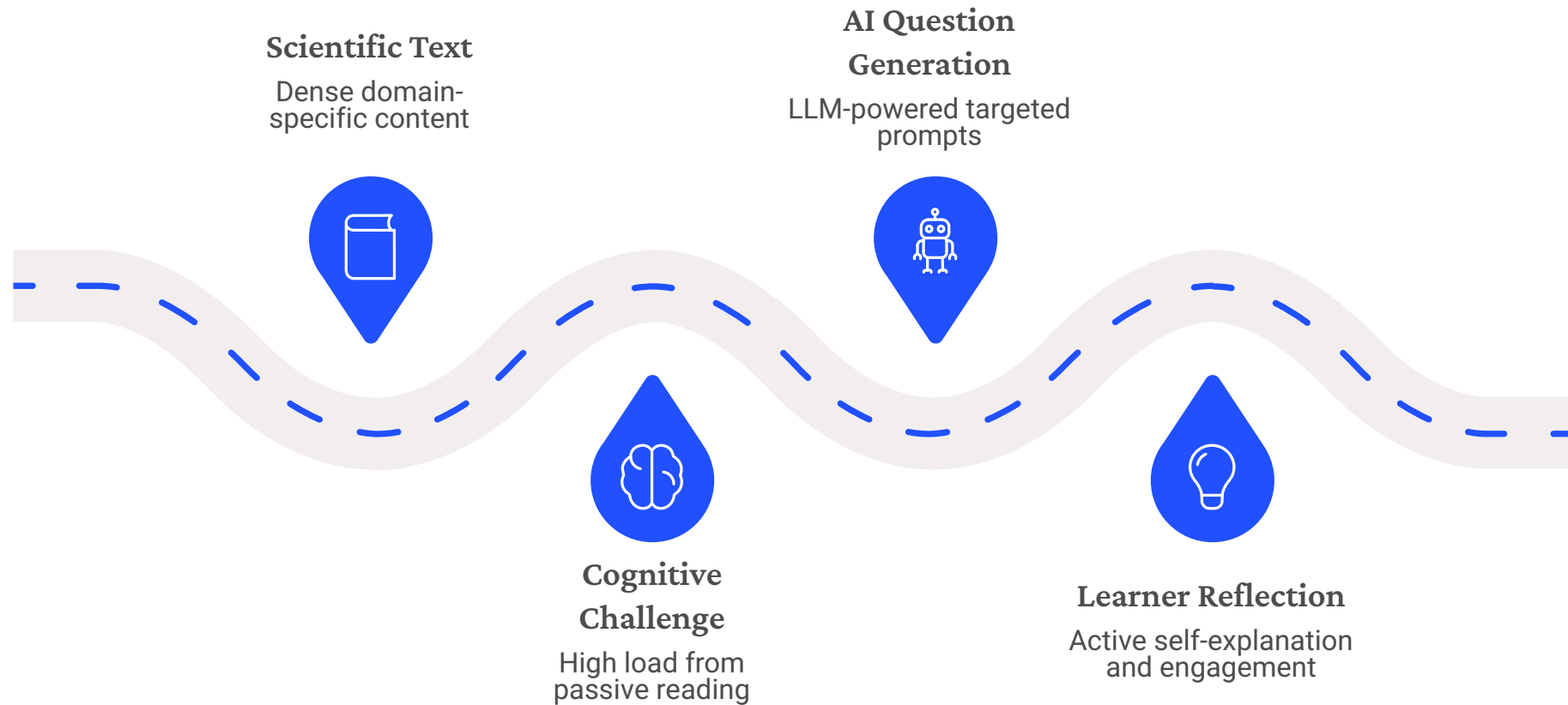


# Research Question

**Can AI-generated questions support reading comprehension and cognitive engagement in learners working with scientific texts?**

This study investigates whether a learning-oriented chatbot, powered by large language models, can scaffold active reading by prompting reflection, self-explanation, and deeper inquiry to support active cognitive engagement during reading.

# From Text to Comprehension



The model positions AI-generated questions as a mediating scaffold — transforming passive exposure into an active, reflective reading experience that supports meaningful comprehension.

# Three Core Research Questions

1

How do students interact with AI based Learning Chatbot?

2

What learning outcomes and cognitive processes are supported?

3

How do motivation and self-efficacy influence chatbot use?

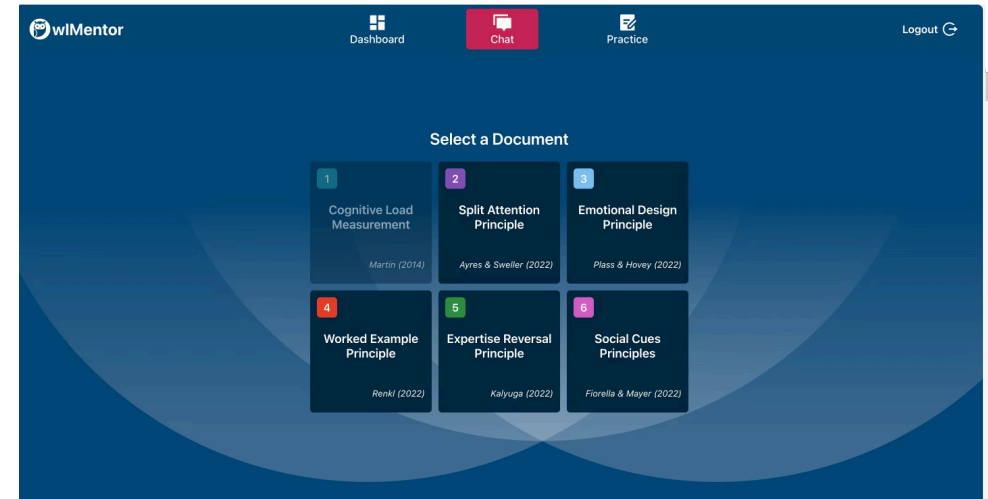
# Introducing OwlMentor

**Definition:** AI-powered GPT-based tool for scientific text comprehension

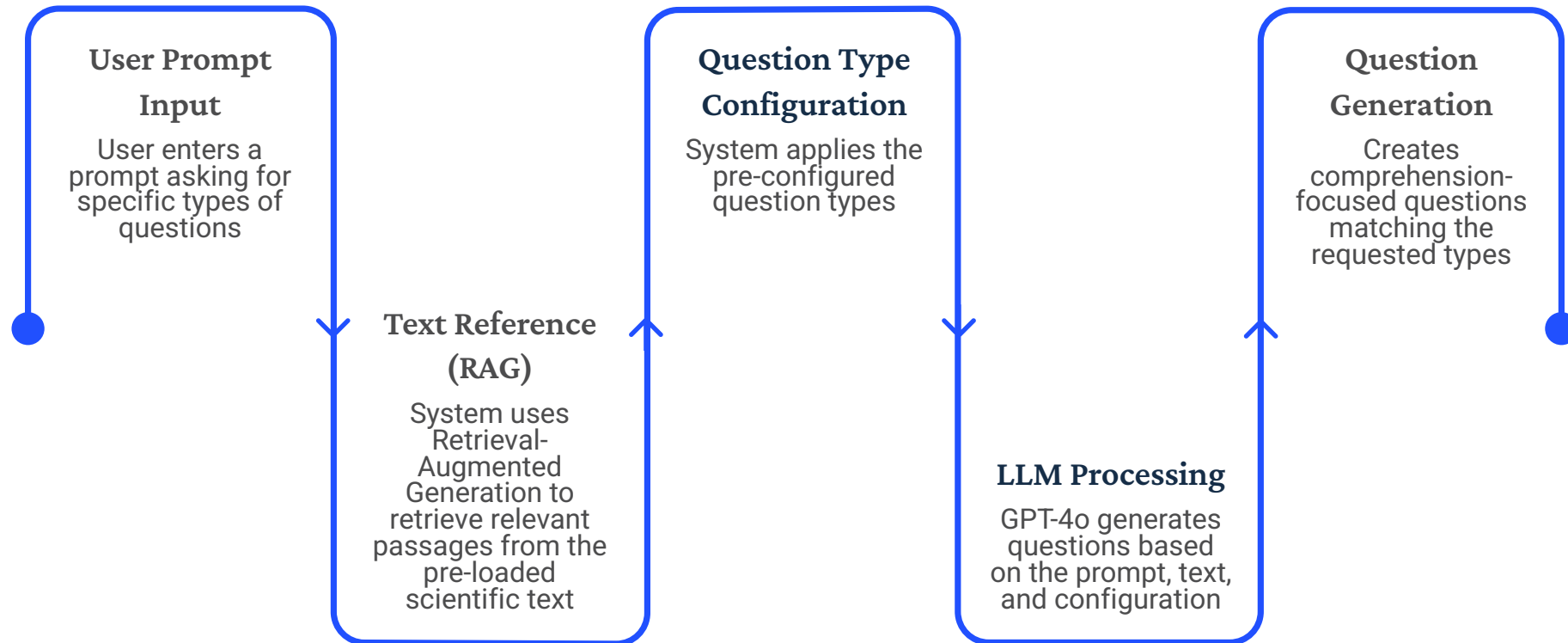
**Technical Approach:** Uses RAG (Retrieval-Augmented Generation) to ground questions in source text

## Key Features:

- Free Chat
- Text Summarization
- Question Generation
- Quiz Creation
- Feedback

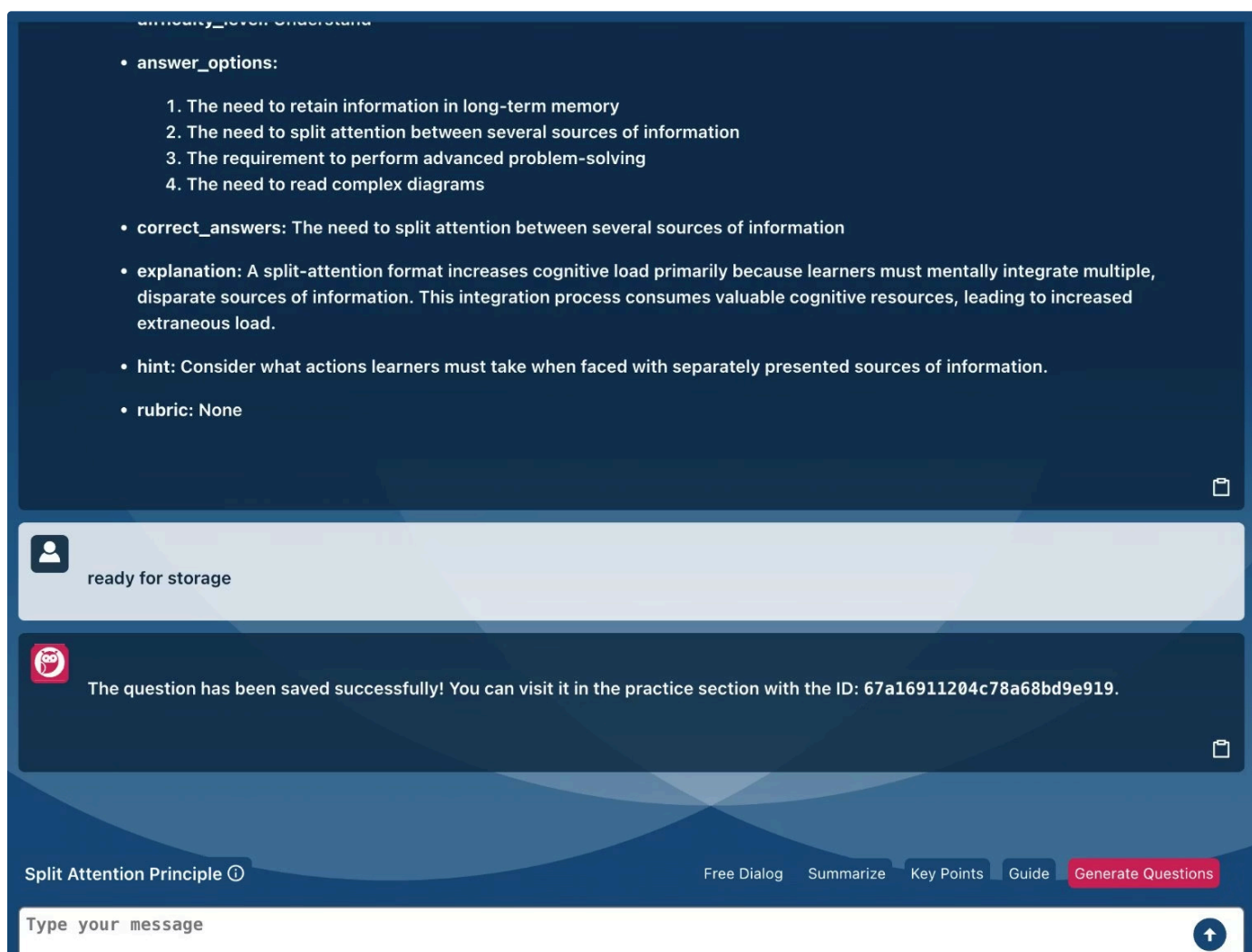
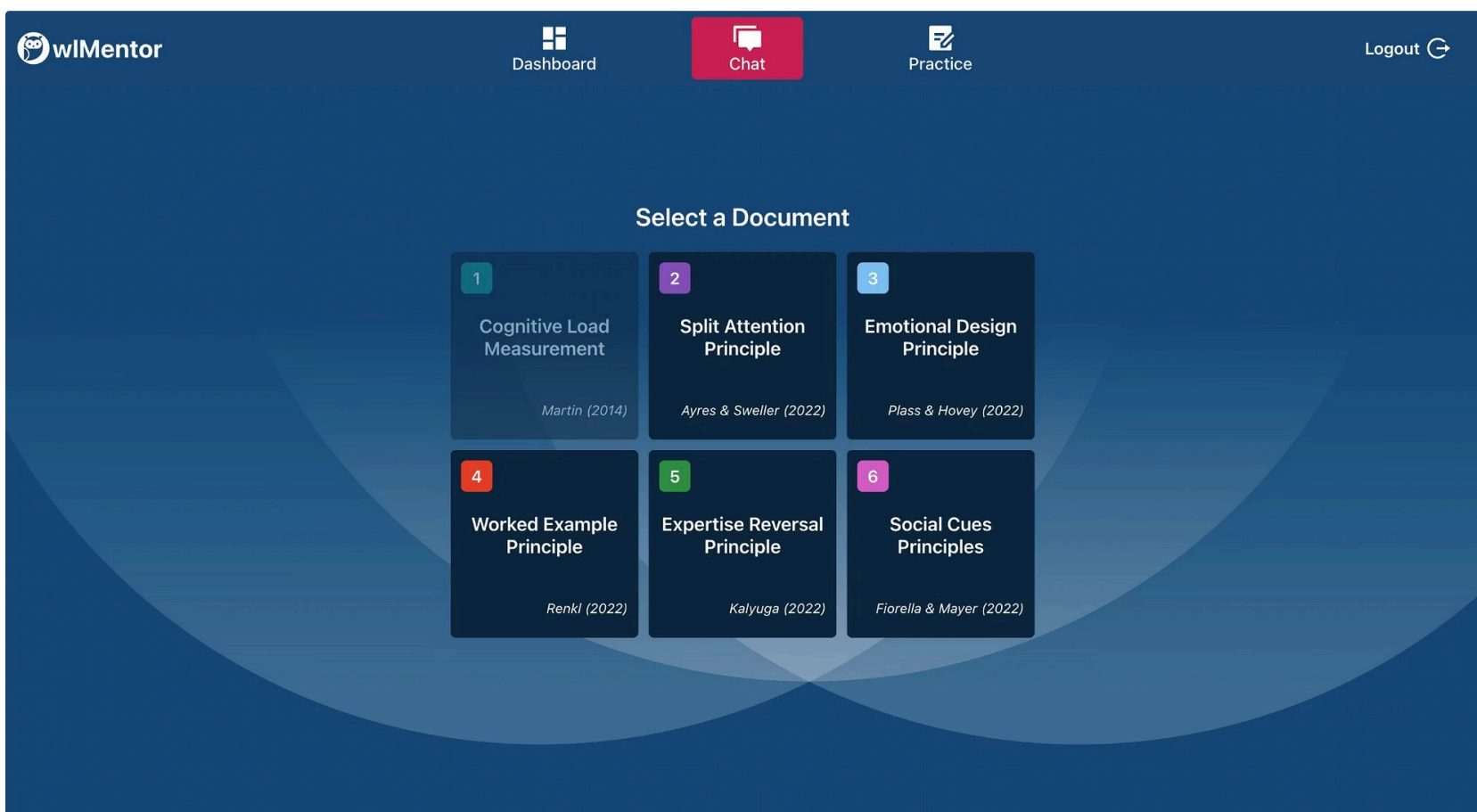
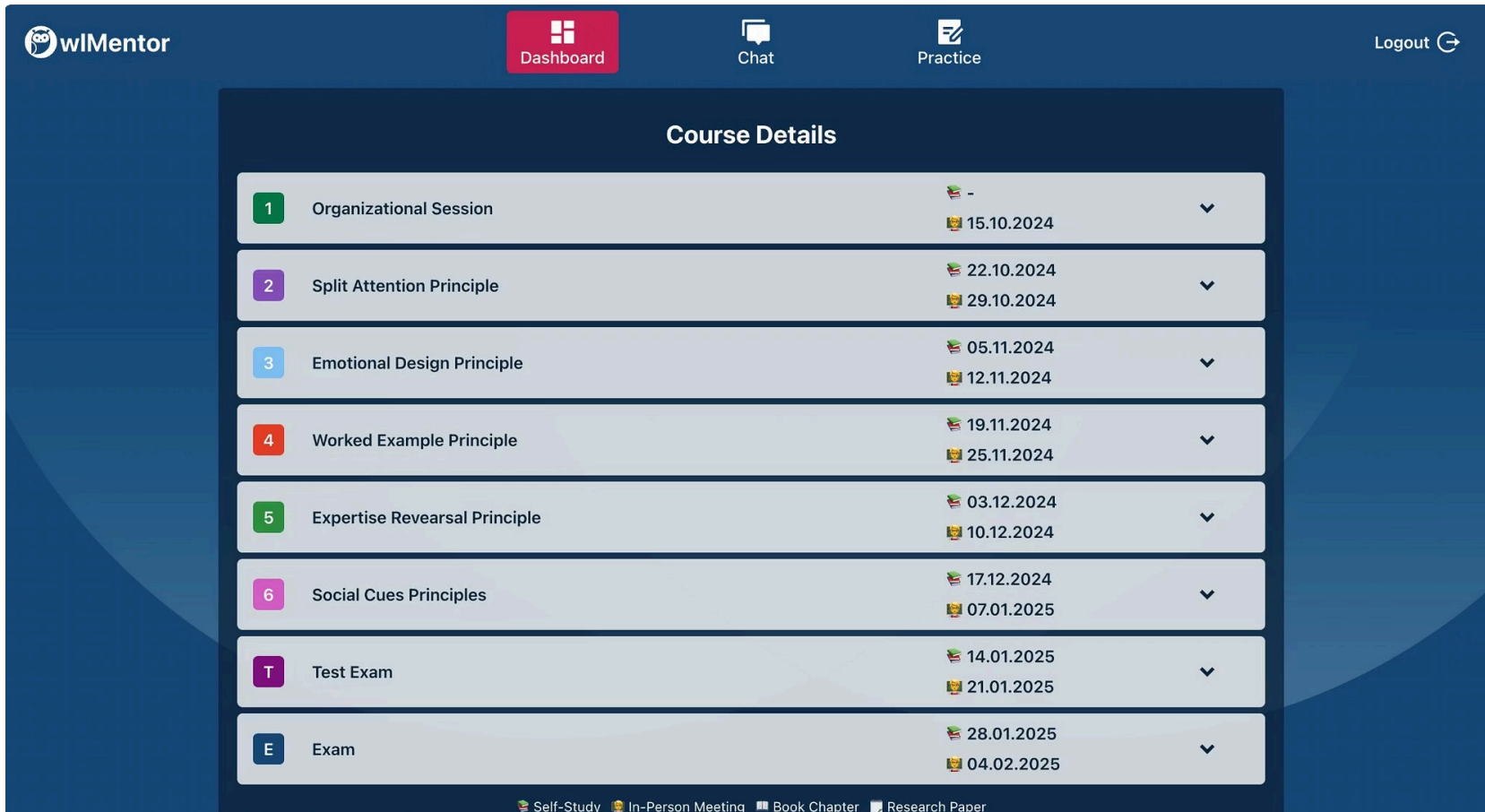


# Technical Pipeline



The workflow starts with a user prompt, not text input. The system then references the pre-loaded scientific text through RAG (Retrieval-Augmented Generation) and the pre-configured question types before using GPT-4o to generate comprehension questions aligned with the requested format.

# AI-Generated Questions in Practice



Generate a [TRUE/FALSE, MULTIPLE CHOICE, or OPEN-ENDED] question for [SECTION OR PAGE] with a difficulty level of [REMEMBER, UNDERSTAND, APPLY, or EVALUATE].

# MCQ



Here are the multiple-choice questions based on Sections 1, 2, and 3, with a focus on the "Remember" difficulty level:

## Section 1: Definition of Split-Attention (Remember)

1. What is a primary characteristic of the split-attention effect?

- Caused by integrating disparate information sources in a single format.
- Involves increased cognitive load due to separating information sources.
- Enhances learning by using multiple simultaneous information sources.
- Leads to reduced errors by keeping diagrammatic information separate.

### Bullet Points:

- **question\_type:** Multiple Choice
- **difficulty\_level:** Remember
- **answer\_options:** Caused by integrating disparate information sources in a single format., Involves increased cognitive load due to separating information sources., Enhances learning by using multiple simultaneous information sources., Leads to reduced errors by keeping diagrammatic information separate.
- **correct\_answers:** Involves increased cognitive load due to separating information sources.
- **explanation:** The split-attention effect is when learners have to split their attention between different information sources, leading to increased cognitive load.
- **hint:** Consider the cognitive load required when information sources are not integrated.

# True/False



## Question: True/False

The split-attention effect is reduced when information is presented in a non-integrated format.

- **question\_type:** True/False
- **difficulty\_level:** Understand
- **answer\_options:** True, False
- **correct\_answers:** False
- **explanation:** The split-attention effect is actually increased when information is presented in a non-integrated format because it raises the extraneous cognitive load by requiring learners to mentally integrate disparate pieces of information.
- **hint:** Consider whether splitting information aids in understanding or whether integrating information supports better comprehension outcomes.

Would you like to refine this question further, or is it ready for storage?



# Open Ended



Generate open ended question for section 1,2,3 with a difficulty level of remember



## Question: Open-Ended

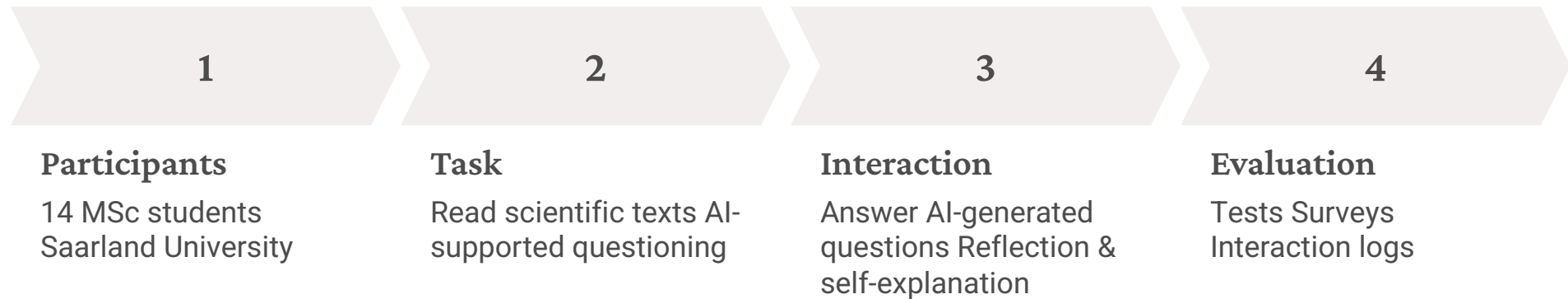
Can you describe the concept of the split-attention effect, providing an example from its historical research origins in instructional design?

- **question\_type:** Open-Ended
- **difficulty\_level:** Remember
- **answer\_options:** None
- **correct\_answers:** The split-attention effect occurs when learners must divide attention between different sources of information, leading to increased cognitive load. Historically, Tarmizi and Sweller's research identified this effect in geometry worked examples where separate diagrams and solutions made learning less effective. Integrated materials, where text and diagrams are combined, improve performance by reducing extraneous cognitive load.
- **explanation:** The question addresses the primary understanding of the split-attention effect and relates it to its initial identification through research that highlights the need for integrating informational sources to improve learning outcomes.
- **hint:** Think about how separating information sources might affect a learner's cognitive resources and how combining those sources could alleviate the issue.
- **rubric:**
  - **3 points:** Clearly describes the split-attention effect and provides a relevant historical example demonstrating understanding of research outcomes.
  - **2 points:** Describes the basic concept but lacks a detailed or accurate example.
  - **1 point:** Mentions the concept with minimal detail or incorrect context.
  - **0 points:** Does not address the split-attention effect or provides irrelevant information.

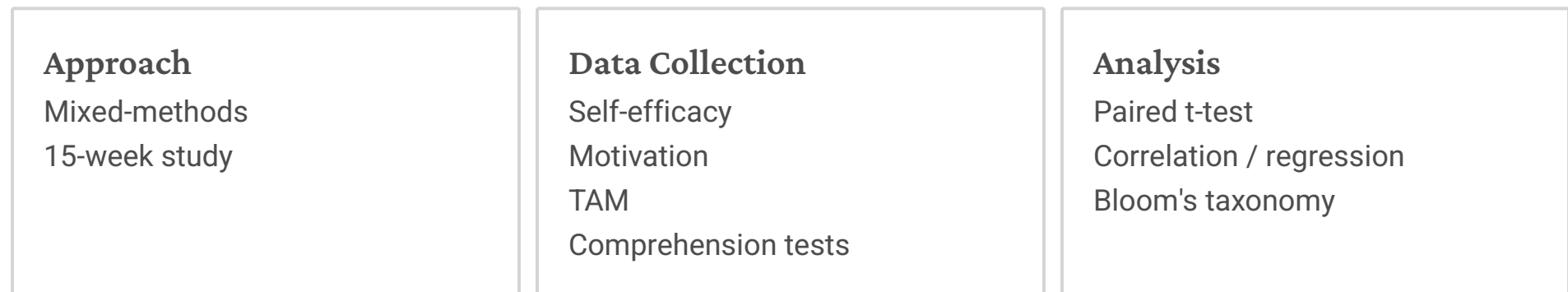
Would you like to refine this question further, or is it ready for storage?

# Study Design & Methodology

## Study Flow



## Methodology

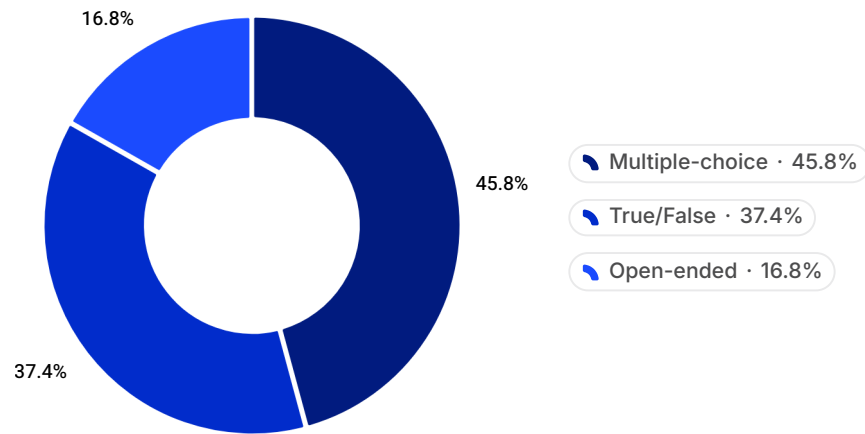


## RESULTS

### RQ1: How do students interact with OwlMentor?

#### Question Generation Patterns

179 total questions generated

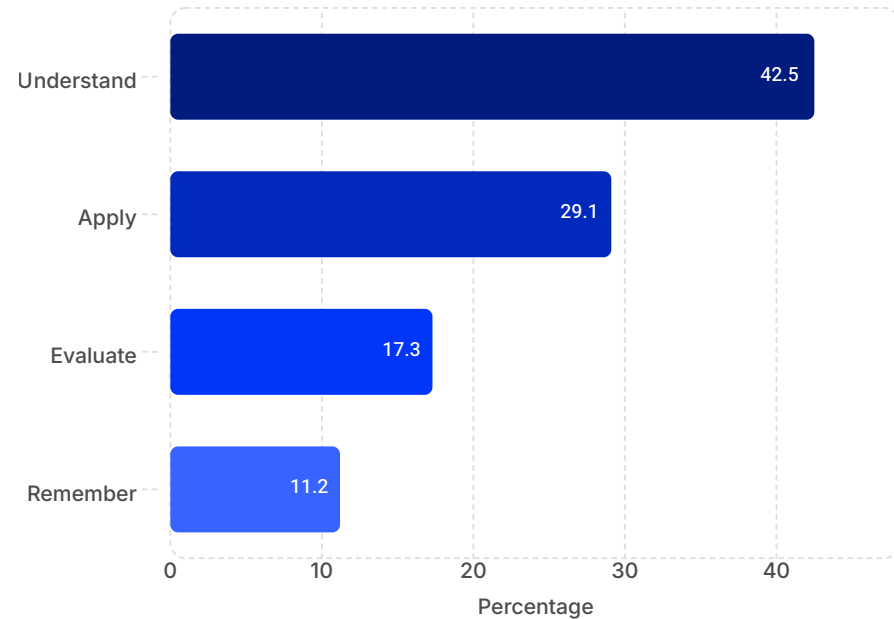


**Key insight:** Multiple-choice questions were the most frequently generated type

#### Cognitive Levels

Bloom's Taxonomy

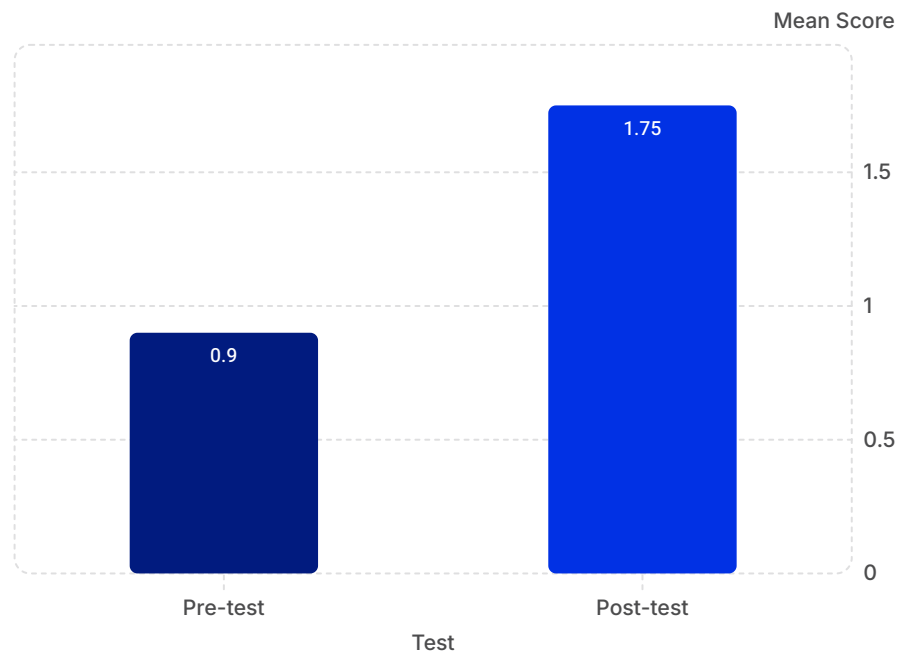
Cognitive Level



**Key insight:** 71.6% of questions targeted mid-level cognitive processes (Understand & Apply)

# RQ2: What learning outcomes and cognitive processes are supported?

### Comprehension Improvement

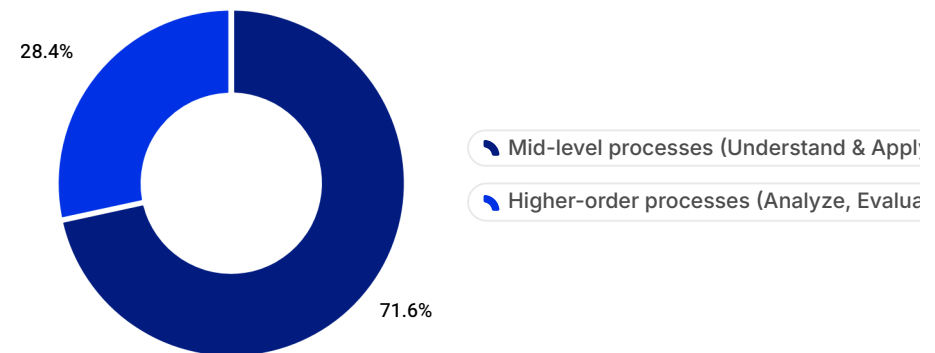


94% improvement ( $p < .001$ ,  $d = 1.78$ )

**Key insight:** Significant learning gains indicate the effectiveness of the system

**Understanding and recall positively correlated** ( $r = 0.66$ ,  $p = .020$ ). This suggests that students who better understood the material were also more likely to retain it.

### Cognitive Process Distribution



**Key insight:** Most questions targeted mid-level cognitive processes

# RQ3: How do motivation and self-efficacy influence chatbot use?

## The AI Motivation Paradox

### Higher Initial AI Motivation

Students with higher initial AI motivation tended to use the chatbot LESS

### Lower Initial AI Motivation

Students with lower initial AI motivation tended to use the chatbot MORE

## Engagement & Usefulness

**Increased Engagement:** Positively related to perceived usefulness and recall

**Motivation & Self-Efficacy remained stable** throughout 15-week intervention

# Practical Implications

## Practical Implications for Educators

- **Target Moderate/Lower Motivation:** OwlMentor is particularly valuable for students with moderate or lower initial motivation
- **Emphasize Perceived Usefulness:** Perceived usefulness drives sustained engagement more than ease of use
- **Elevate Prompt Engineering:** Prompt engineering should encourage higher-order thinking (Analyze, Evaluate, Create)
- **Integrate with Course Structure:** Integration into course structure (graded activities) increases consistent use
- **AI as Collaborative Partner:** AI tools work best as collaborative partners, not passive content sources

# Limitations & Future Research Directions

From present constraints to the next phase of evidence-building.

## Study Limitations

### Small sample size

N=14 with gender imbalance (11 female, 3 male)

### Single institution

Saarland University - limited generalizability

### Self-report bias

Motivation and engagement measures may be influenced by self-report bias

### Single evaluator

Comprehension assessments lacked inter-rater reliability

### No control group

There was no comparison group for benchmarking outcomes

### Short duration

One semester was insufficient for long-term effects

# Future Research Directions

## Study Design Improvements

- Larger and more diverse samples
- Controlled experimental designs

## Long-Term Learning Effects

- Longitudinal studies
- Delayed post-tests
- Transfer of learning

## Evaluation & Comparison

- Multi-evaluator assessments
- Comparative studies across AI tools

## Intelligent System Design

- Adaptive personalization



## CONCLUSION

# AI as a Cognitive Scaffold for Learning

### Enhanced Comprehension

The system demonstrates substantial gains in understanding complex texts.

### Active Engagement

OwlMentor shifts students from passive content consumption to active knowledge construction.

### Scaffolds Critical Thinking

The system effectively cultivates essential self-questioning skills for deeper learning.

### Practical Application

These findings highlight the immense potential for AI in educational practices.

# References

A selection of academic sources that underpin the theoretical framework and methodologies discussed in this presentation.

- **Bhat, S., Nguyen, H. A., Moore, S., Stamper, J., & Sakr, M. (2022).** Towards automated generation and evaluation of questions in educational domains. *Proceedings of the 15th International Conference on Educational Data Mining*, 701–704.
- **Black, P., & Wiliam, D. (1998).** Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- **Bruner, J. S. (1966).** *Toward a theory of instruction*. Harvard University Press.
- **Chin, C., & Brown, D. E. (2000).** Learning in science: A comparison of deep and surface approaches. *Journal of Research in Science Teaching*, 37(2), 109–138.
- **Davis, F. D. (1989).** Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- **Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989).** User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003.
- **Deci, E. L., & Ryan, R. M. (2000).** The "what" and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268.
- **Dijkstra, R., Genç, Z., Kayal, S., & Kamps, J. (2022).** Reading comprehension quiz generation using generative pretrained transformers. *iTextbooks @ AIED 2022*, XXX-XXX.
- **Heilman, M., & Smith, N. A. (2010).** Good question! Statistical ranking for question generation. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 609–617). <https://aclanthology.org/N10-1086/>
- **Holmes, W., Bialik, M., & Fadel, C. (2019).** *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- **Hsu, C.-L., & Chiu, C.-M. (2004).** Internet self-efficacy and electronic service acceptance. *Decision Support Systems*, 38(3), 369–381.
- **Oche, A. J., Folashade, A. G., Ghosal, T., & Biswas, A. (2025).** A systematic review of key retrieval-augmented generation (RAG) systems: Progress, gaps, and future directions. *arXiv preprint*. <https://arxiv.org/abs/2507.18910>
- **Park, S. Y. (2009).** An analysis of the technology acceptance model in understanding e-learning usage behavior. *Journal of Educational Technology & Society*, 12(3), 150–162.
- **Pintrich, P. R., & De Groot, E. V. (1990).** Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33–40.
- **Rosenshine, B., Meister, C., & Chapman, S. (1996).** Teaching students to generate questions: A review of intervention studies. *Review of Educational Research*, 66(2), 181–221.
- **Siriwardhana, S., Weerasekera, R., Wen, E., Kaluarachchi, T., Rana, R., & Nanayakkara, S. (2023).** Improving the domain adaptation of retrieval-augmented generation (RAG) models for open-domain question answering. *Transactions of the Association for Computational Linguistics*, 11, 1–17. <https://doi.org/10.1162/tacla00530>
- **Snow, C. E. (2002).** *Reading for understanding: Toward an R&D program in reading comprehension*. RAND Corporation.
- **Thüs, D., Malone, S., & Brünken, R. (2024).** Student AI question co-creation for enhancing reading comprehension. *IEEE Transactions on Learning Technologies*. Advance online publication. <https://doi.org/10.1109/TLT.2023.3333439>
- **Türkyilmaz, İ., Yıldırım, S., & Akçay, D. (2023).** Student attitudes toward AI tutors: Role of curiosity and perceived utility. *Computers & Education*, 189, 104598.
- **Venkatesh, V., & Bala, H. (2008).** Technology Acceptance Model 3 and a research agenda on interventions. *Decision Sciences*, 39(2), 273–315.
- **Venkatesh, V., & Davis, F. D. (2000).** A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, 46(2), 186–204.

# Thank You

Piumal Mahawasala

M.Sc. Educational Technology

Saarland University

[kyma00002@uni-saarland.de](mailto:kyma00002@uni-saarland.de)

[piumal.mahawasala@dfki.de](mailto:piumal.mahawasala@dfki.de)

## Questions?

