National Consortium for Languages Education

LanguageHubs

A Framework for Teacher-AI Complementarity

Professor Mutlu Cukurova

University College London

m.cukurova@ucl.ac.uk

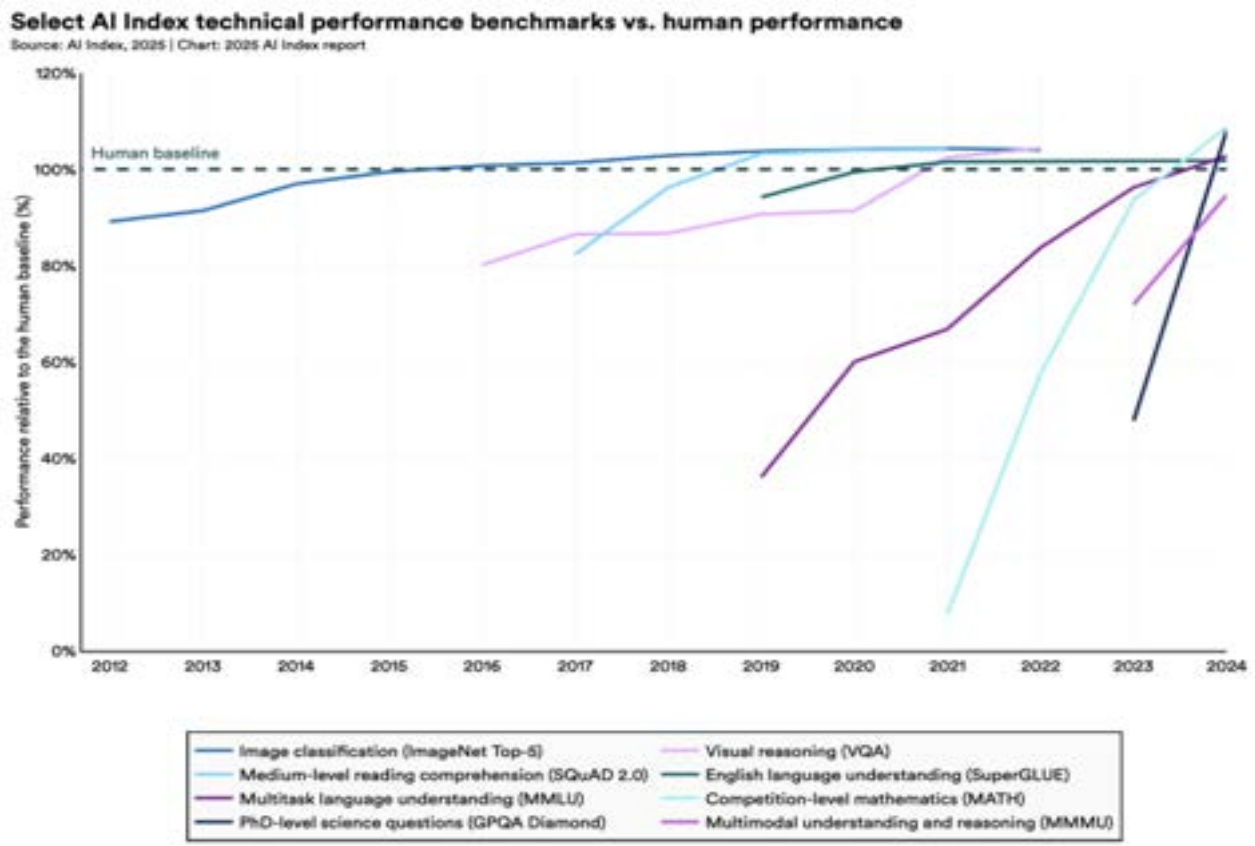
@mutlucukurova

Knowledge Lab UCL UCLAIT - Team

Taico Teacher-AI Complementarity

UCL

AI performance continues to improve



The 2025 AI Index Report (2025). Human-centered Artificial Intelligence Centre, Stanford University.

Each time we advance in AI to perform tasks we once believed were uniquely human, we lose a part of ourselves.

A fundamental question to ask is …

What is the core of a human that we can not cut away anymore?

What should we educate people about?

What should the role of an AI system be in education?

How should we design education to cultivate the distinctly human capacities that AI cannot replicate?

AI Implications for Education

1. Educating people about AI so that they can use it effectively and ethically

2. Design and development of AI to support teaching and learning processes

3. Innovation in education to sustain the distinctly human capacities & prepare people for an AI-ubiquitous world

Luckin, R., & Cukurova, M. (2019). Designing Educational Technologies in the Age of AI: A Learning Sciences Driven Approach. *British Journal of Educational Technology*, *50*(6), 490-504.

Three Conceptualisations of AI in Education

AI can be conceptualised to externalize, be internalized or extend human cognition.

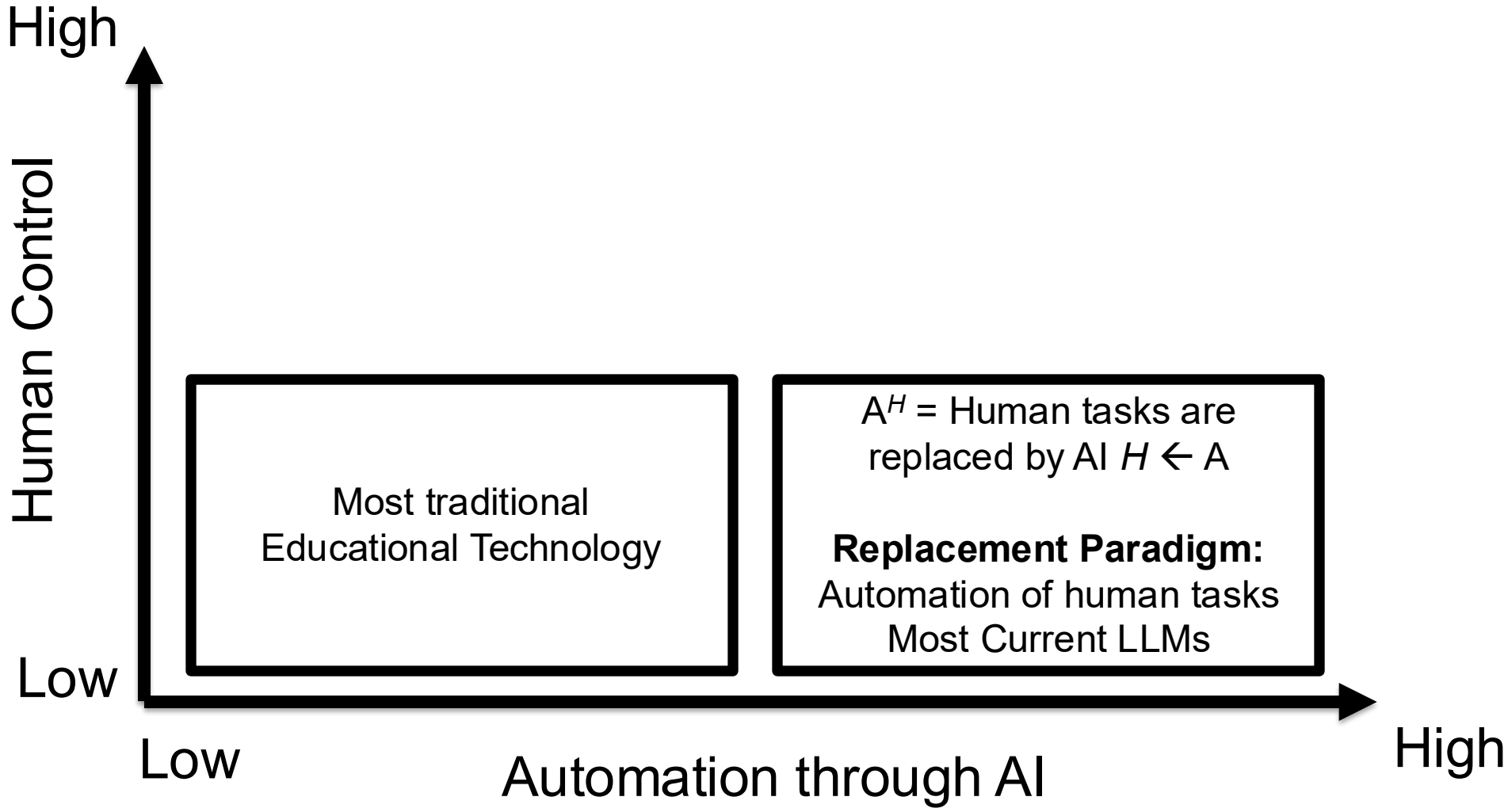
- A^H = Human tasks are replaced by AI H ← A
- H^A = Humans can internalise AI models H → A
- H[A] = Human (H) extended with an AI (A), **tightly coupled synergistic human and AI systems.**

$$H[A] \neq H + A \quad \rightarrow \quad H[A] > {}_{max}(H, A)$$

The whole should be more than the sum of its parts.
Human + AI is better than the max of human or AI alone.
Changes in H, also in A, are expected.

Cukurova, M. (2019). Learning Analytics as AI Extenders in Education: Multimodal Machine Learning versus Multimodal Learning Analytics. *Proceedings of the Artificial Intelligence and Adaptive Education Conference, xx1-xx3.*
Cukurova, M. (2024). The Interplay of Learning, Analytics, and Artificial Intelligence in Education. *British Journal of Educational Technology*.
https://doi.org/10.48550/arXiv.2403.16081

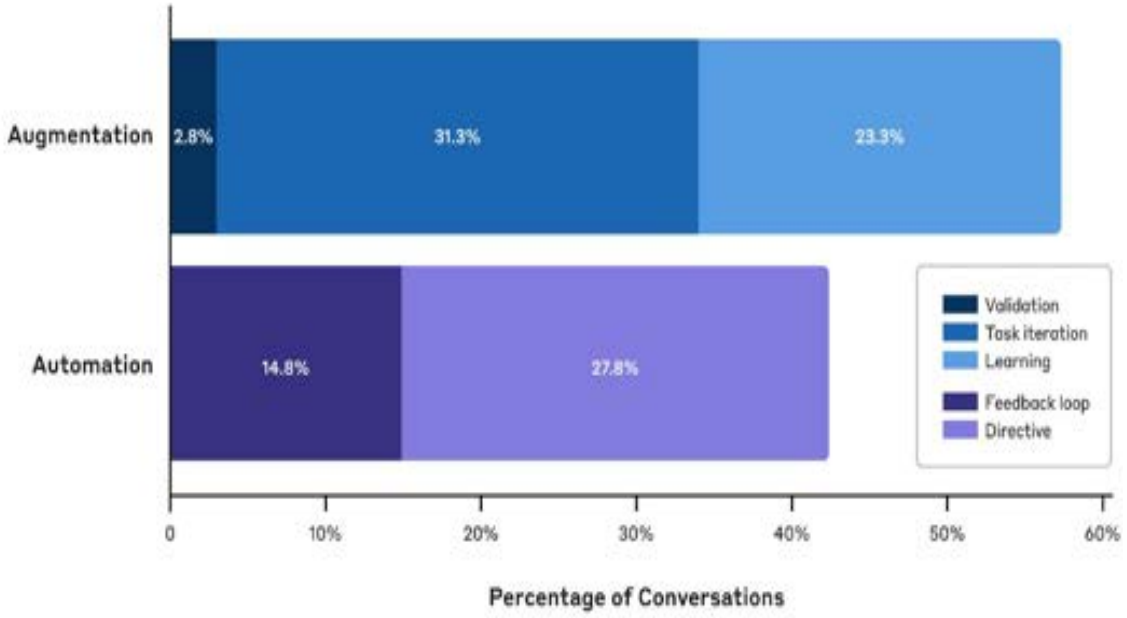
AI in Education: A vision for the future



Cukurova, M. (2024). The Interplay of Learning, Analytics, and Artificial Intelligence in Education. *British Journal of Educational Technology*. https://doi.org/10.48550/arXiv.2403.16081

How exactly genAI is used in practice?

- Based on four million Claude.ai conversations, 4% of occupations show usage for **at least** 75% of their professional tasks.
- e.g. **Foreign Language & Literature Teachers**: AI usage for planning course content, generating teaching materials, not for maintaining student records.



Automative Behaviors *AI directly executes tasks with minimal human involvement*	**Augmentative Behaviors** *AI enhances human capabilities through collaboration*
Directive: Complete task delegation with minimal interaction *Illustrative Example: "Format this technical documentation in Markdown"*	**Task Iteration:** Collaborative refinement process *Illustrative Example: "Let's draft a marketing strategy for our new product. ... Good start, but can we add some concrete metrics?"*
Feedback Loop: Task completion guided by environmental feedback *Illustrative Example: "Here's my Python script for data analysis – it's giving an IndexError. Can you help fix it? ... Now I'm getting a different error..."*	**Learning:** Knowledge acquisition and understanding *Illustrative Example: "Can you explain how neural networks work?"*
	Validation: Work verification and improvement *Illustrative Example: "I've written this SQL query to find duplicate customer records. Can you check if my logic is correct and suggest any improvements?"*

Handa et al., (2025). Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations. Anthropic, Technical Report.

What is the impact of genAI on teachers?

- GenAI-assisted lesson and resource preparation on teacher time *vs* approaches unassisted by genAI.
- 68 representative schools across the UK, 259 KS3 Science Teachers, an extensive range of teaching experience.
- Planning time for GenAI teachers was 56.2 minutes per week compared to 81.5 minutes in the non-GenAI group, **a reduction of 31% in preparation time for teachers.**
- **No statistically significant difference in the quality of resources.**

Roy, P., Poet, H., Staunton, R., Aston, K., & Thomas, D. (2024). *ChatGPT in lesson preparation: A teacher choices trial evaluation report*. EEF

AI-generated Feedback

Assignment

Q8. Now, categorise children scoring below the 15th percentile as 'Low'. You can create a new variable to indicate if the child is above or under the 15th percentile.

Student response

percentile_cutoff <- qnorm(0.15) data$low_parent <- ifelse(data$ptotal_z < percentile_cutoff, "Low", "Not Low") data$low_teacher <- ifelse(data$ttotal_z < percentile_cutoff, "Low", "Not Low") head(data[, c("ptotal_z", "ttotal_z", "low_parent", "low_teacher")])

Feedback variants

Feedback 1

Great job on creating the new variables! Your approach using qnorm(0.15) to find the cutoff is correct and shows good understanding. However, consider using "1" and "0" instead of "Low" and "Not Low" for easier numerical analysis. Also, remember to explain your code briefly to demonstrate your reasoning.

Feedback 2

Excellent job, smart to save the percentile! Good job on that and on the labeling using ifelse. I would encourage you to use the tidy syntax to make the steps here clearer to follow, but well done!

Feedback 3

Excellent job! Your approach using qnorm() is precise and efficient. The ifelse() statements are well-constructed for categorization. Consider using the tidyverse syntax for improved readability, but overall, your solution is effective and demonstrates a solid understanding of the concept.

Feedback variants re-evaluation

Here is where those feedback come from :

Feedback 1 was generated by ***AI***

Feedback 2 was generated by ***TAs***

Feedback 3 was generated by ***AI and TAs (co-produced)***

Please take a moment to read again those feedbacks :

Feedback 1
AI

Great job on creating the new variables! Your approach using qnorm(0.15) to find the cutoff is correct and shows good understanding. However, consider using "1" and "0" instead of "Low" and "Not Low" for easier numerical analysis. Also, remember to explain your code briefly to demonstrate your reasoning.

You can now change your responses. Taking into account the source of the feedback (AI), to what extent do you now associate Feedback 1 above with the following terms?

Feedback 2
TAs

Excellent job, smart to save the percentile! Good job on that and on the labeling using ifelse. I would encourage you to use the tidy syntax to make the steps here clearer to follow, but well done!

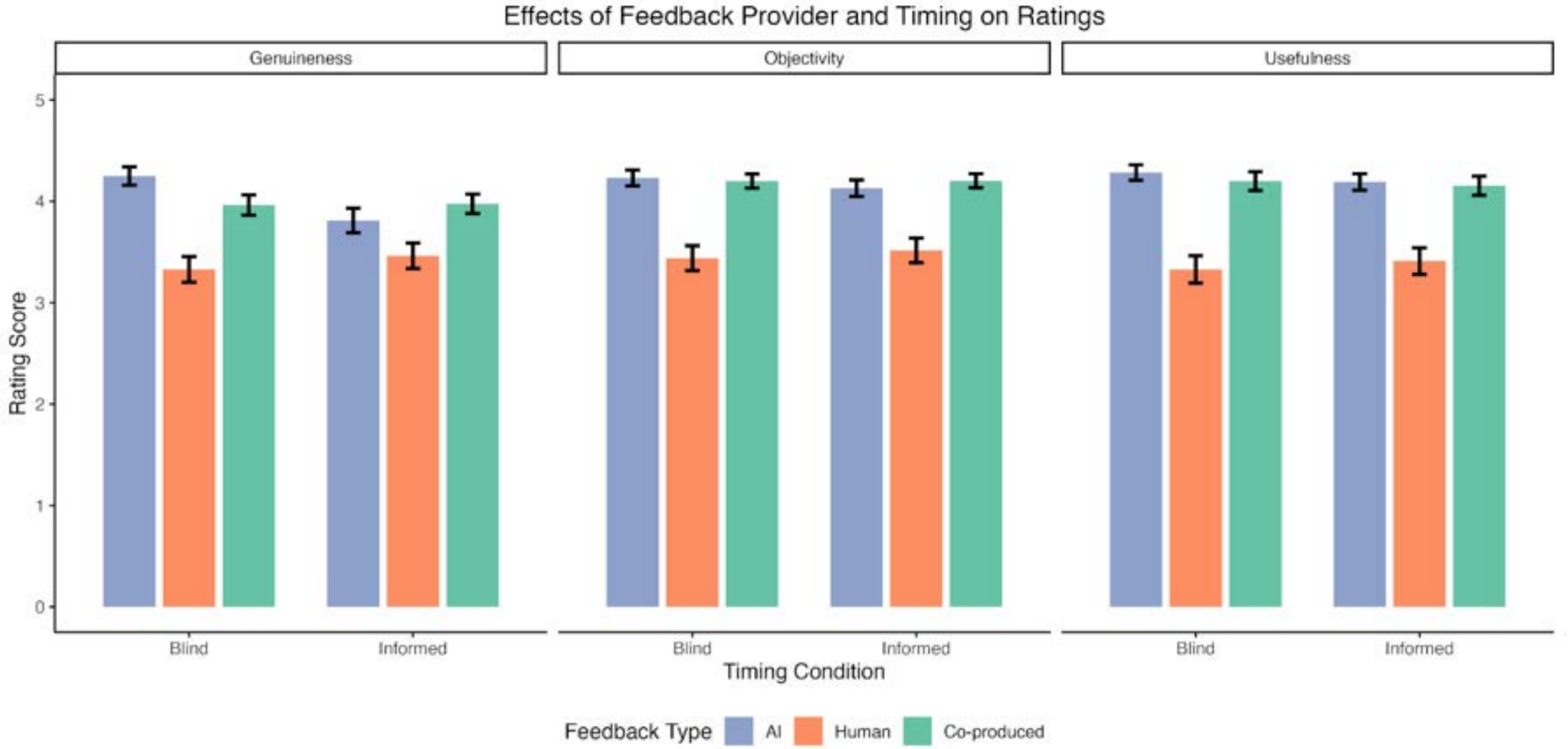
You can now change your responses. Taking into account the source of the feedback (TAs), to what extent do you now associate Feedback 2 above with the following terms?

Feedback 3
AI and TAs (co-produced)

Excellent job! Your approach using qnorm() is precise and efficient. The ifelse() statements are well-constructed for categorization. Consider using the tidyverse syntax for improved readability, but overall, your solution is effective and demonstrates a solid understanding of the concept.

You can now change your responses. Taking into account the source of the feedback (AI and TAs (co-produced)), to what extent do you now associate Feedback 3 above with the following terms?

Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T., & Cukurova, M. (2025). Evaluating Trust in AI, Human, and Co-produced Feedback Among Undergraduate Students. https://arxiv.org/abs/2504.10961



Zhang, A., Gao, Y., Suraworachet, W., Nazaretsky, T., & Cukurova, M. (2025). Evaluating Trust in AI, Human, and Co-produced Feedback Among Undergraduate Students . https://arxiv.org/abs/2504.10961 .

Replacement paradigm can provide productivity gains in certain tasks, but quality improvements at scale need alternative conceptualisations.

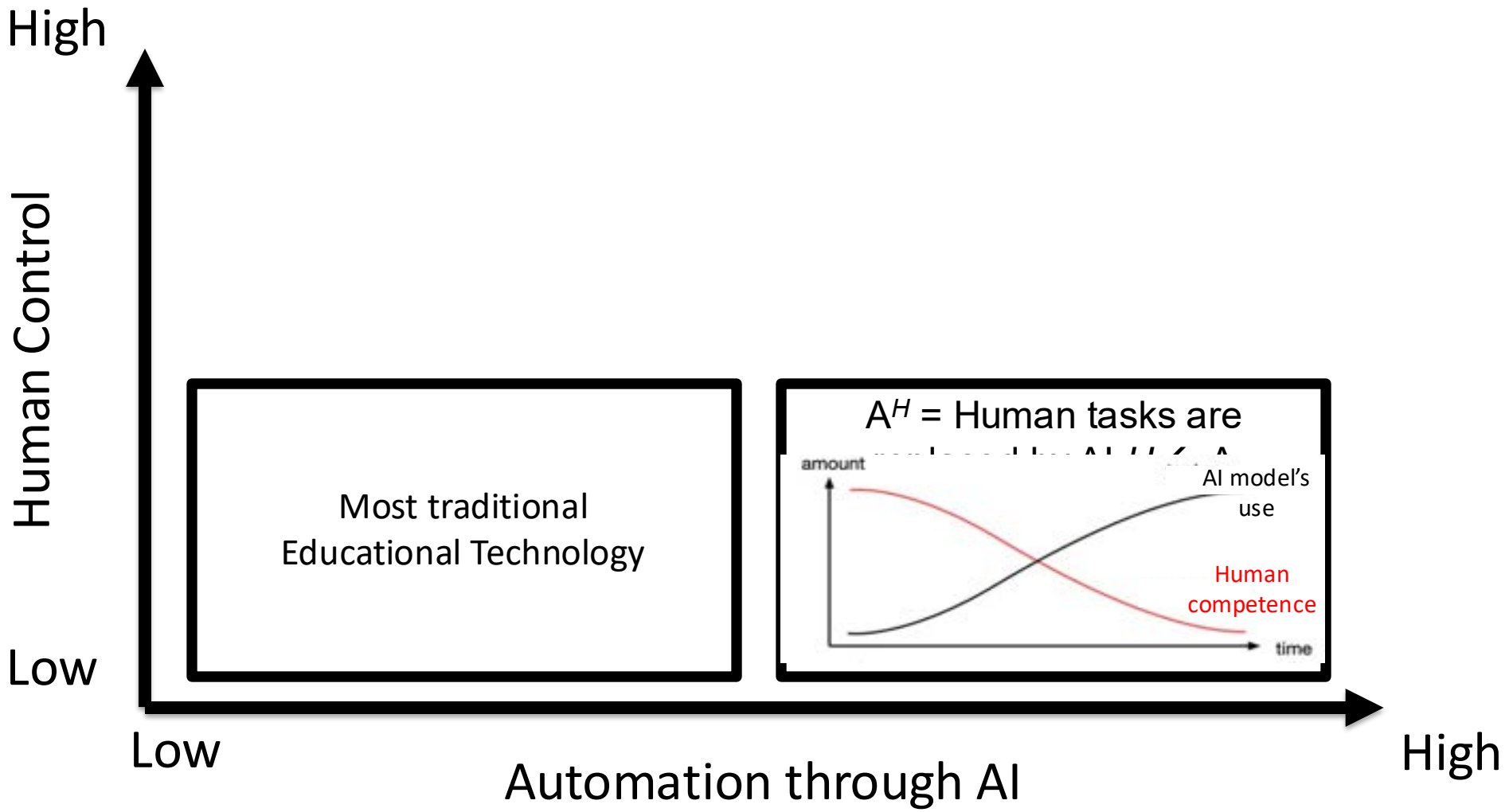
Many other challenges with the replacement paradigm

- Accuracy & Reliability
- Bias
- Digital divide / Equity
- Standardisation of Education
- De-humanising Education
- Cognitive atrophy
- Socio-cultural issues regarding the commercial AI use

Cukurova, M. (2025). Promoting and Protecting Teacher Agency in the Age of Artificial Intelligence. UNESCO Publishing

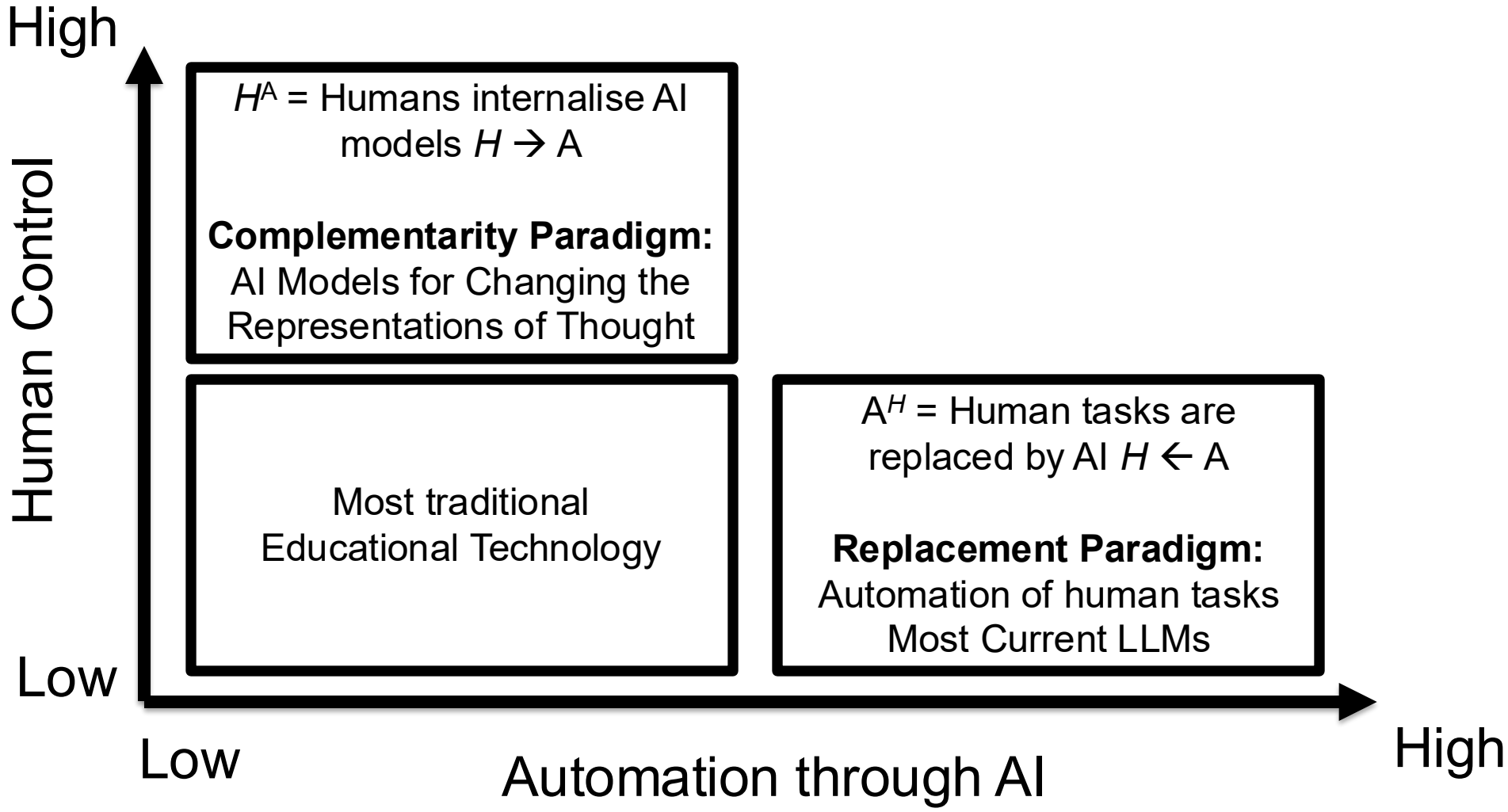
Productivity gains from the replacement paradigm will come at a cost, we need informed decisions about when (or if) it is acceptable.

AI in Education: A vision for the future



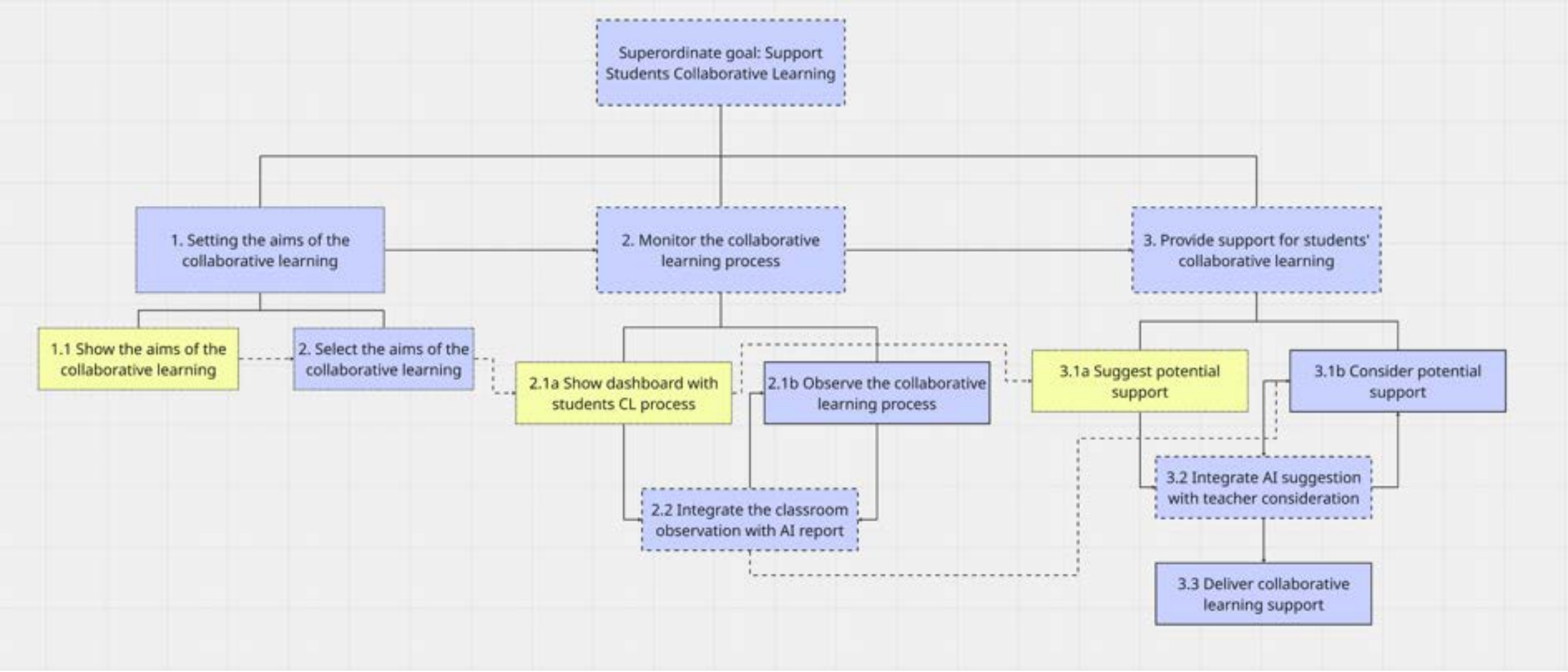
Cukurova, M. (2024). The Interplay of Learning, Analytics, and Artificial Intelligence in Education. *British Journal of Educational Technology*. https://doi.org/10.48550/arXiv.2403.16081

AI in Education: A vision for the future

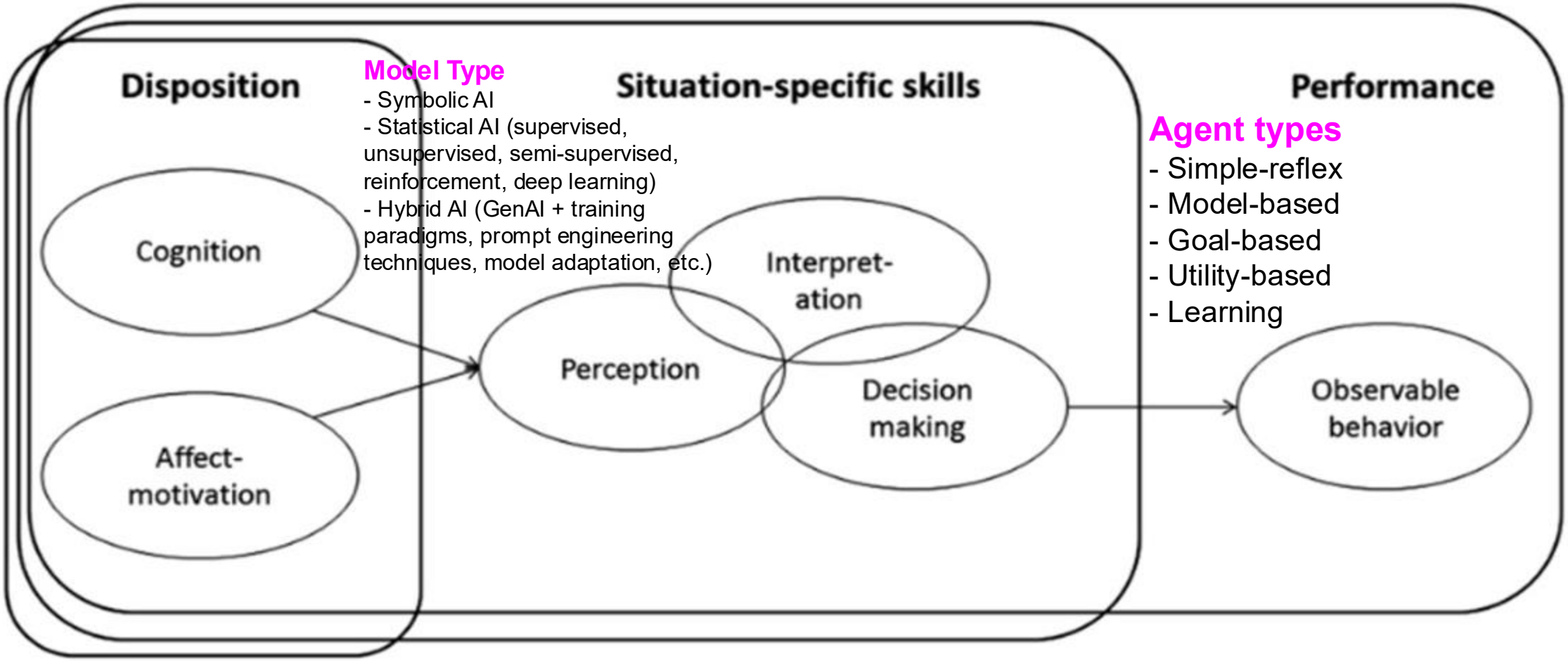


Cukurova, M. (2024). The Interplay of Learning, Analytics, and Artificial Intelligence in Education. *British Journal of Educational Technology*. https://doi.org/10.48550/arXiv.2403.16081

Hierarchical Task Analysis

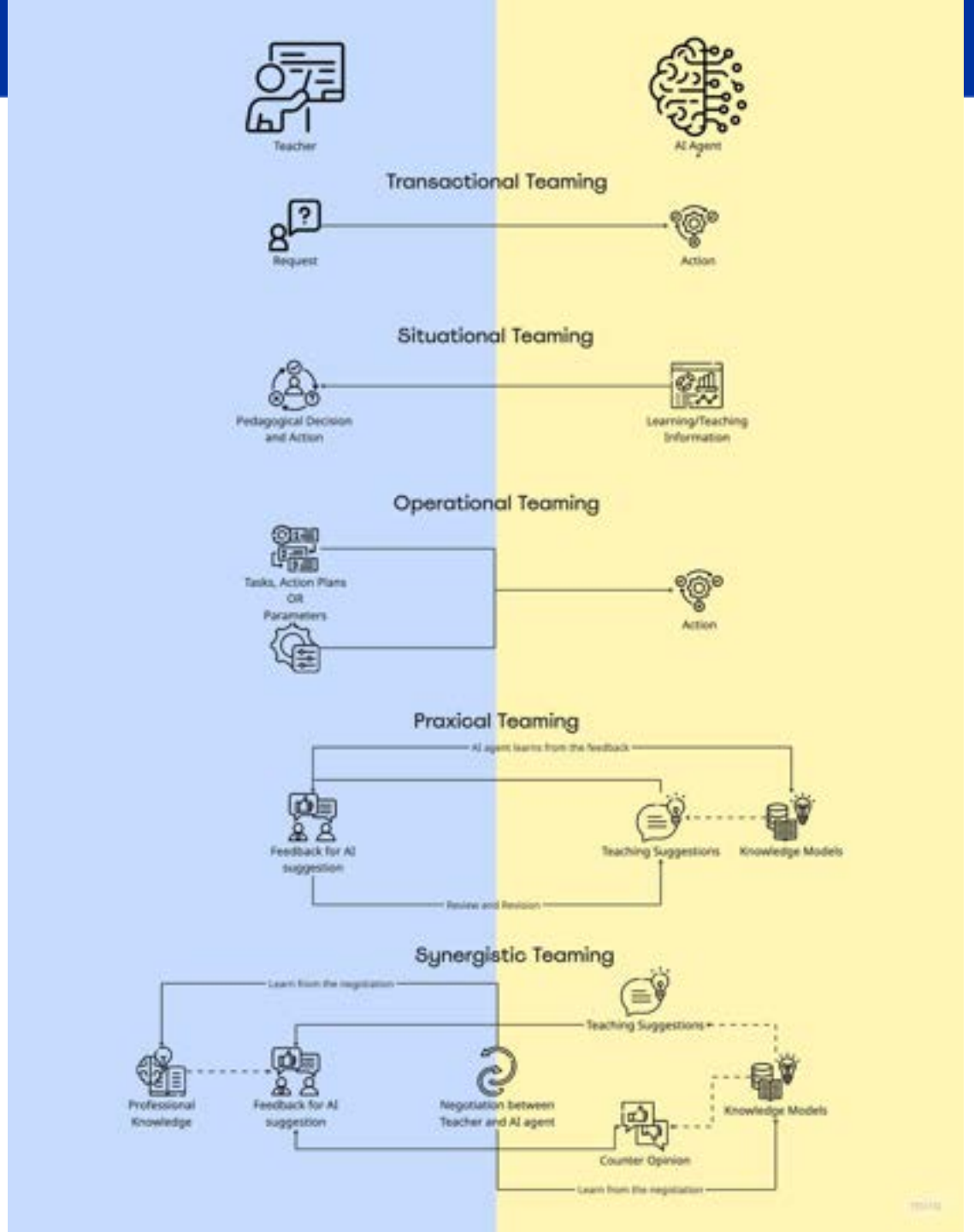


Human competence as a continuum



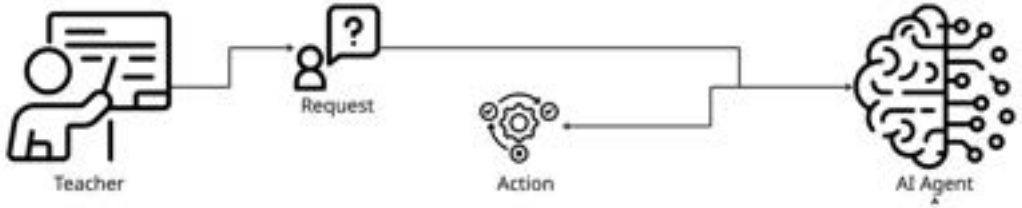
Blömeke et al., (2015) Competence as a continuum

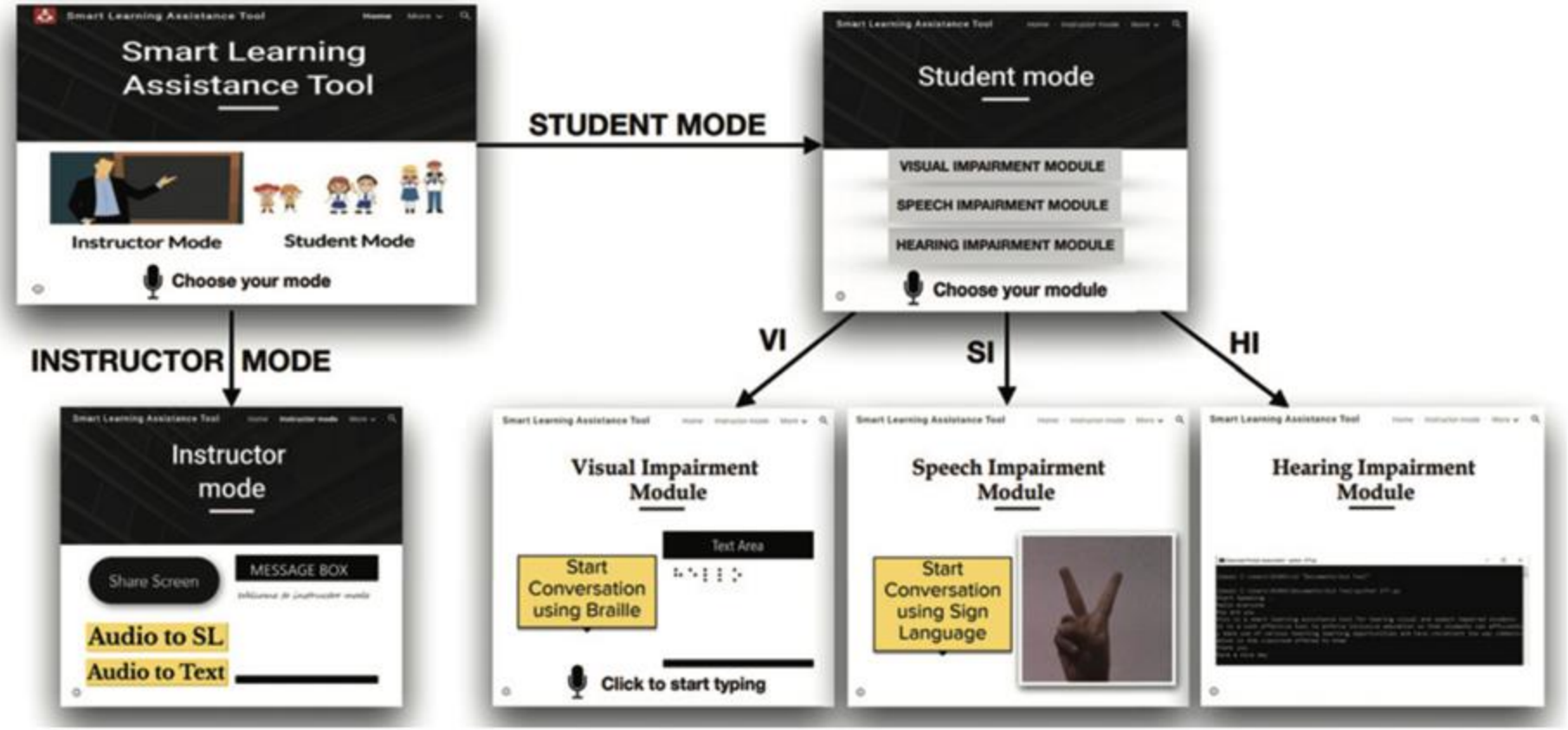
For any given task, depending on the specific teachers' competence and specific AI's affordances



Cukurova, M. (2026) A Conceptual Framework for Teacher-AI Teaming in Education: Harnessing GenAI to Enhance Teacher Agency. OECD Digital Education Outlook.

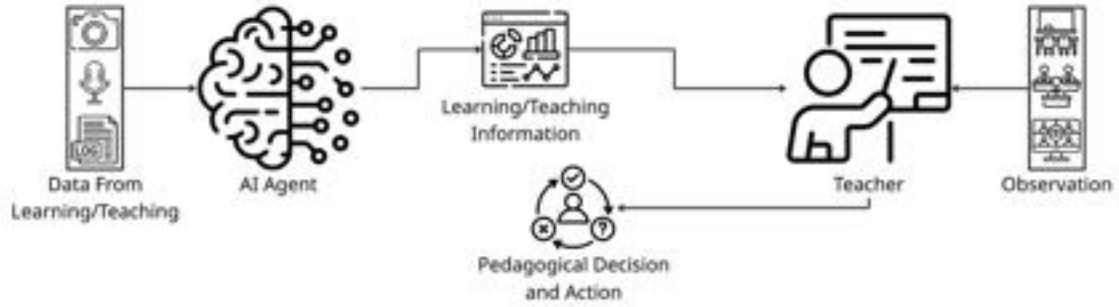
Transactional Teaming

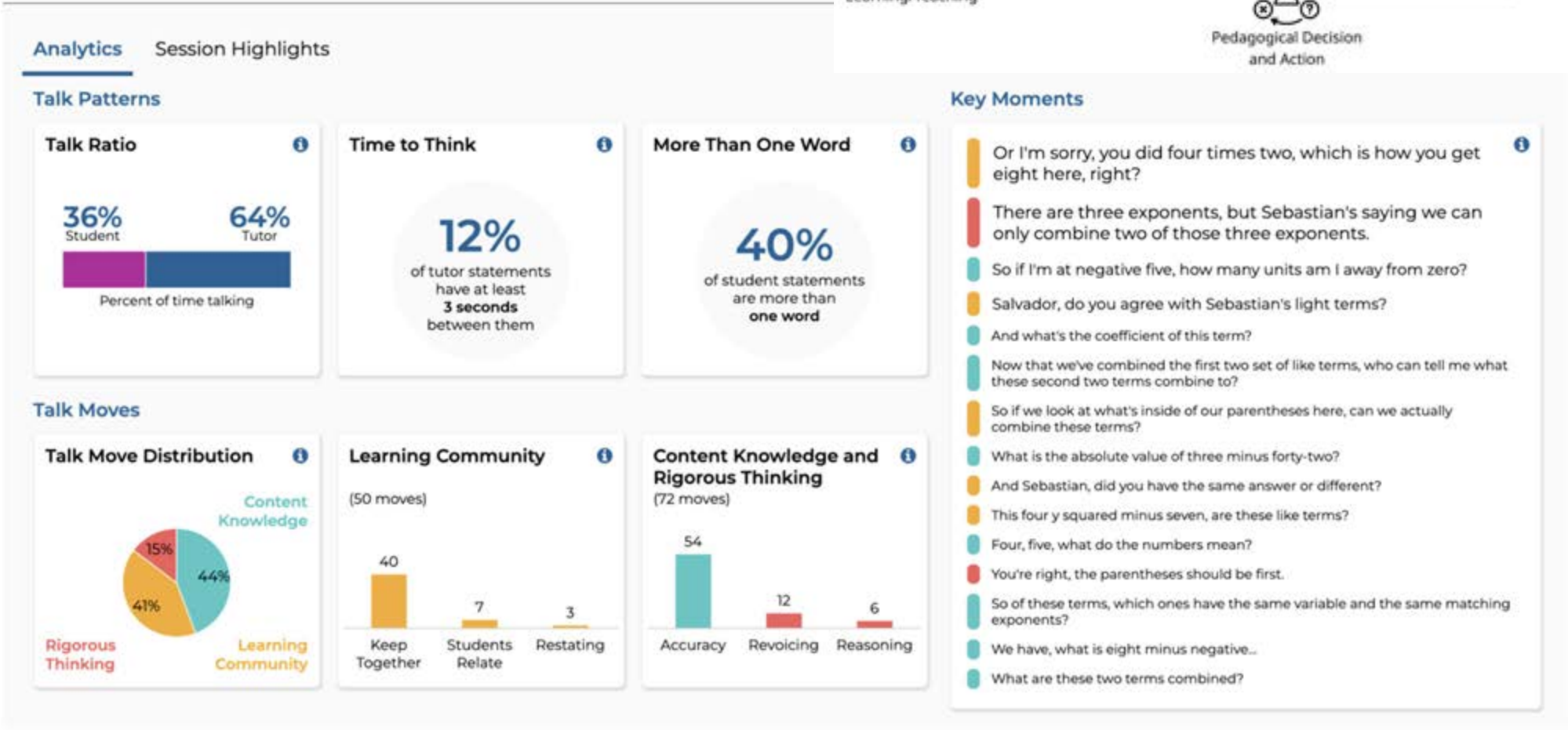




Srivastava, S., Varshney, A., Katyal, S., Kaur, R., & Gaur, V. (2021). A smart learning assistance tool for inclusive education. *Journal of Intelligent & Fuzzy Systems*, *40*(6), 11981–11994. https://doi.org/10.3233/JIFS-210075

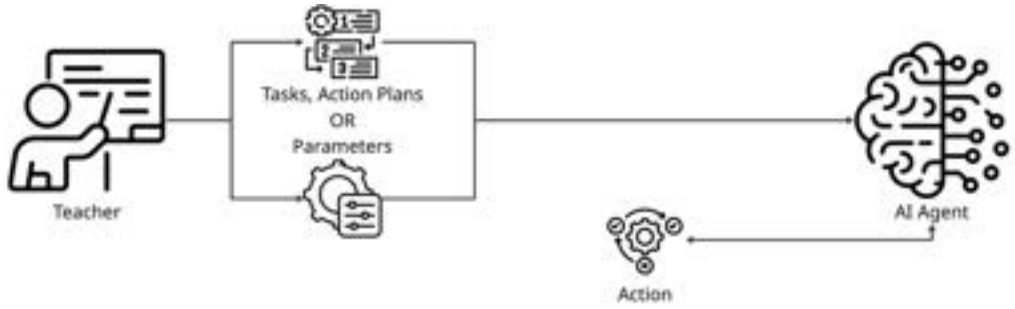
Situational Teaming

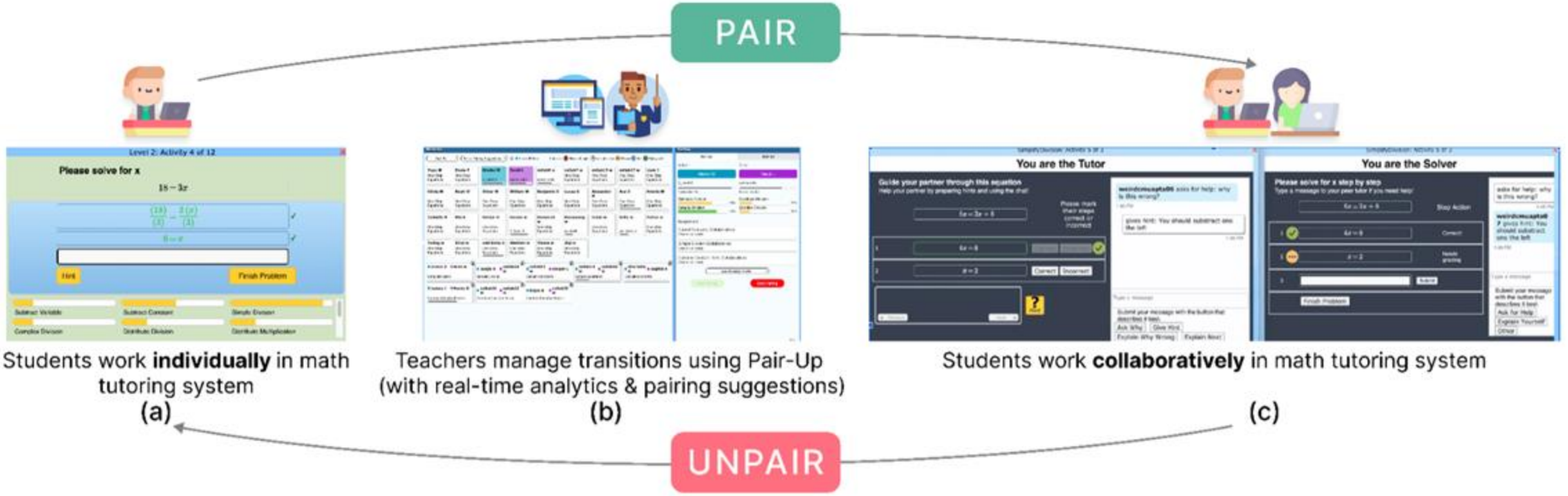




Sawaya, S., Jacobs, J., Moulder, R., Chandler, C., Milne, B., Fischaber, T., & D'Mello, S. K. (2025). Improving Tutor Discourse Practices via AI-Enhanced Coaching: A Piecewise Latent Growth Curve Modeling Approach. In A. I. Cristea, E. Walker, Y. Lu, O. C. Santos, & S. Isotani (Eds), *Artificial Intelligence in Education* (Vol. 15880, pp. 76–89). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-98459-4_6

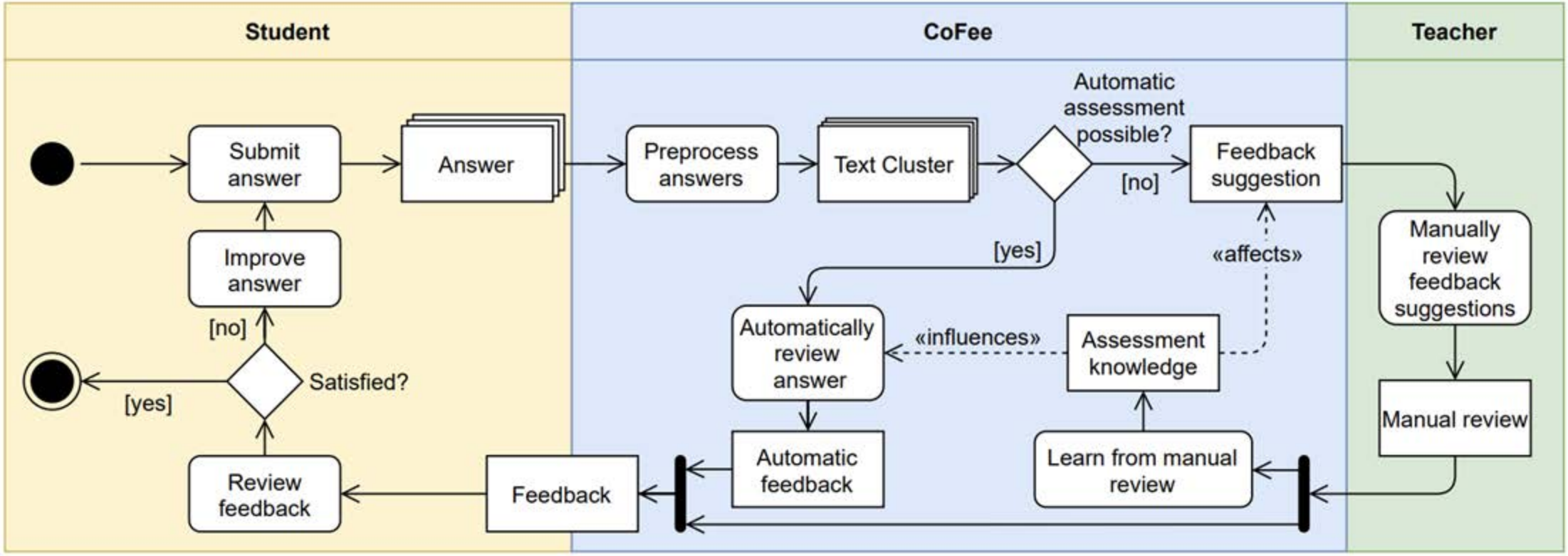
Operational Teaming





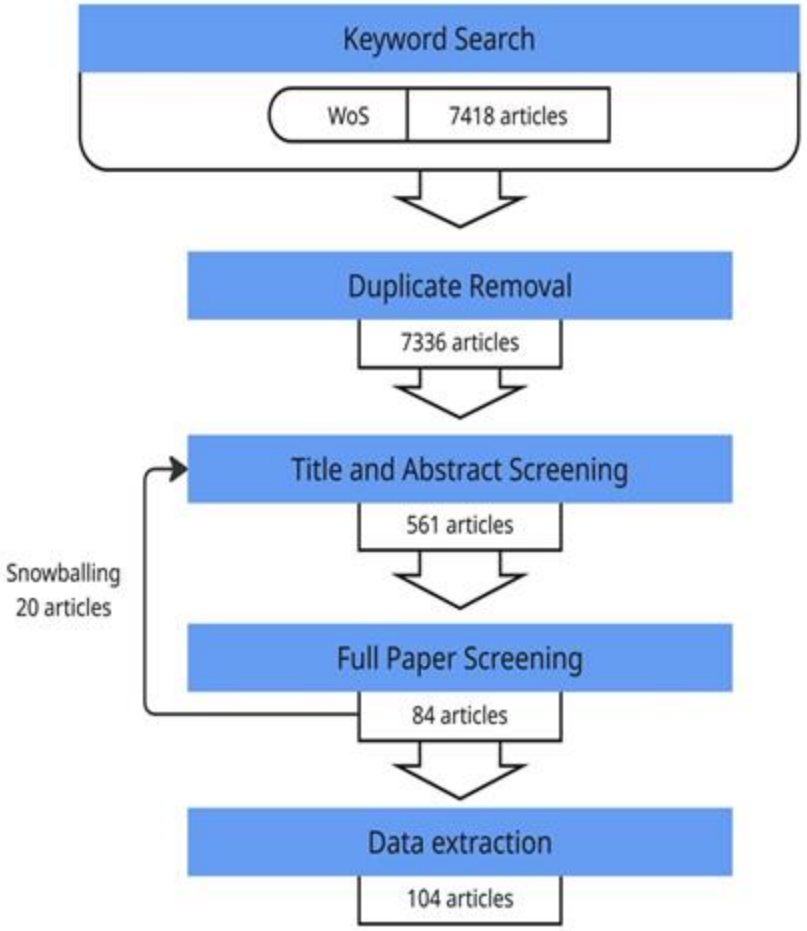
Yang, K. B., Echeverria, V., Lu, Z., Mao, H., Holstein, K., Rummel, N., & Aleven, V. (2023). Pair-Up: Prototyping Human-AI Co-orchestration of Dynamic Transitions between Individual and Collaborative Learning in the Classroom. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3544548.3581398

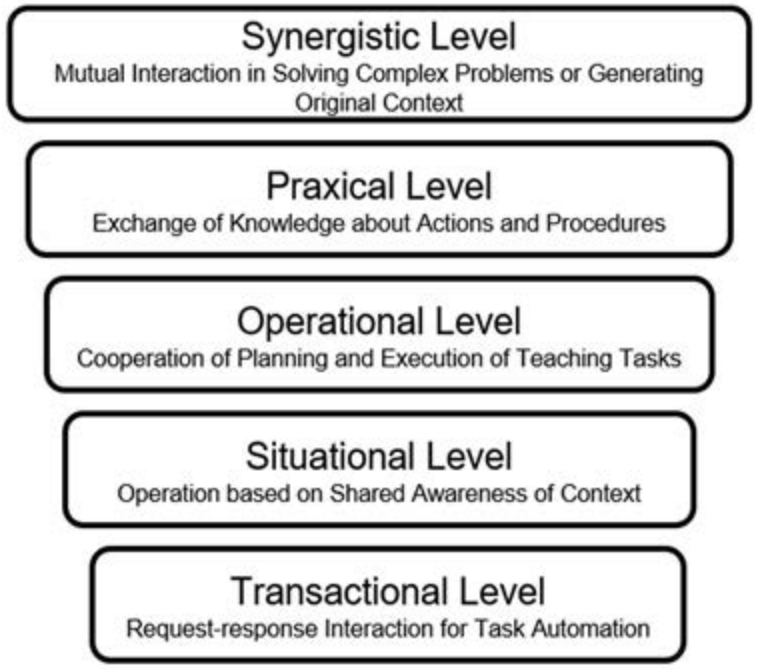
Praxical Teaming

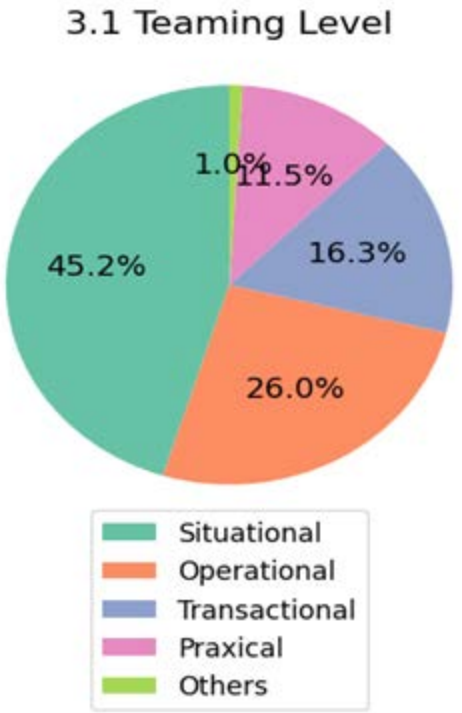


Bernius, J. P., Krusche, S., & Bruegge, B. (2021). A Machine Learning Approach for Suggesting Feedback in Textual Exercises in Large Courses. *Proceedings of the Eighth ACM Conference on Learning @ Scale*, 173–182. https://doi.org/10.1145/3430895.3460135

SLR of Teacher-facing AI – No Synergy





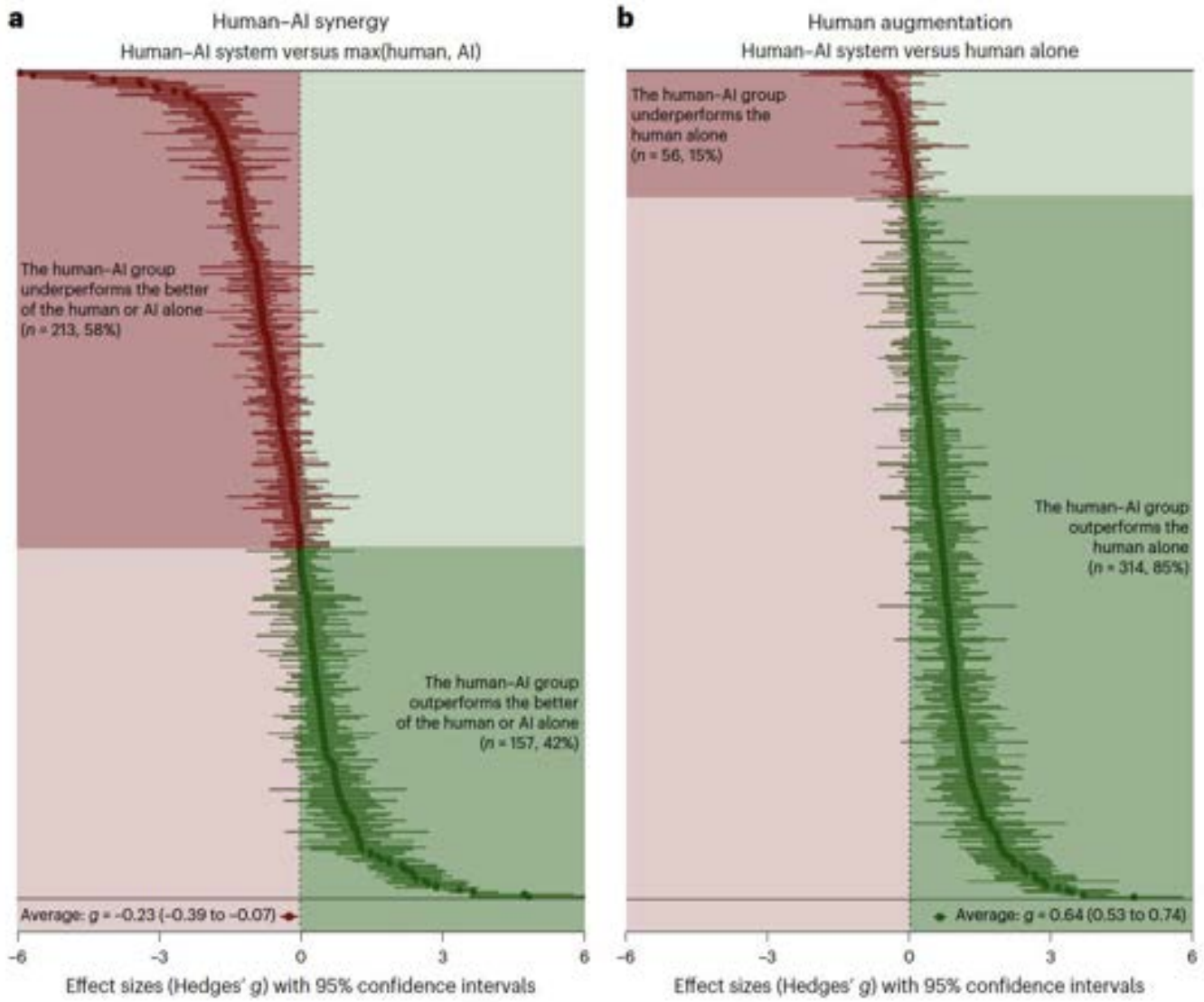


Cukurova, M., Suraworachet, W., Zhou, Q, & Bulathwela, S. (2026). Towards Synergistic Teacher-AI Interactions with Generative Artificial Intelligence. *The handbook of generative AI in Education: Integrating research into practice.* Palgrave Macmillan

Synergistic Human-AI teaming is Difficult

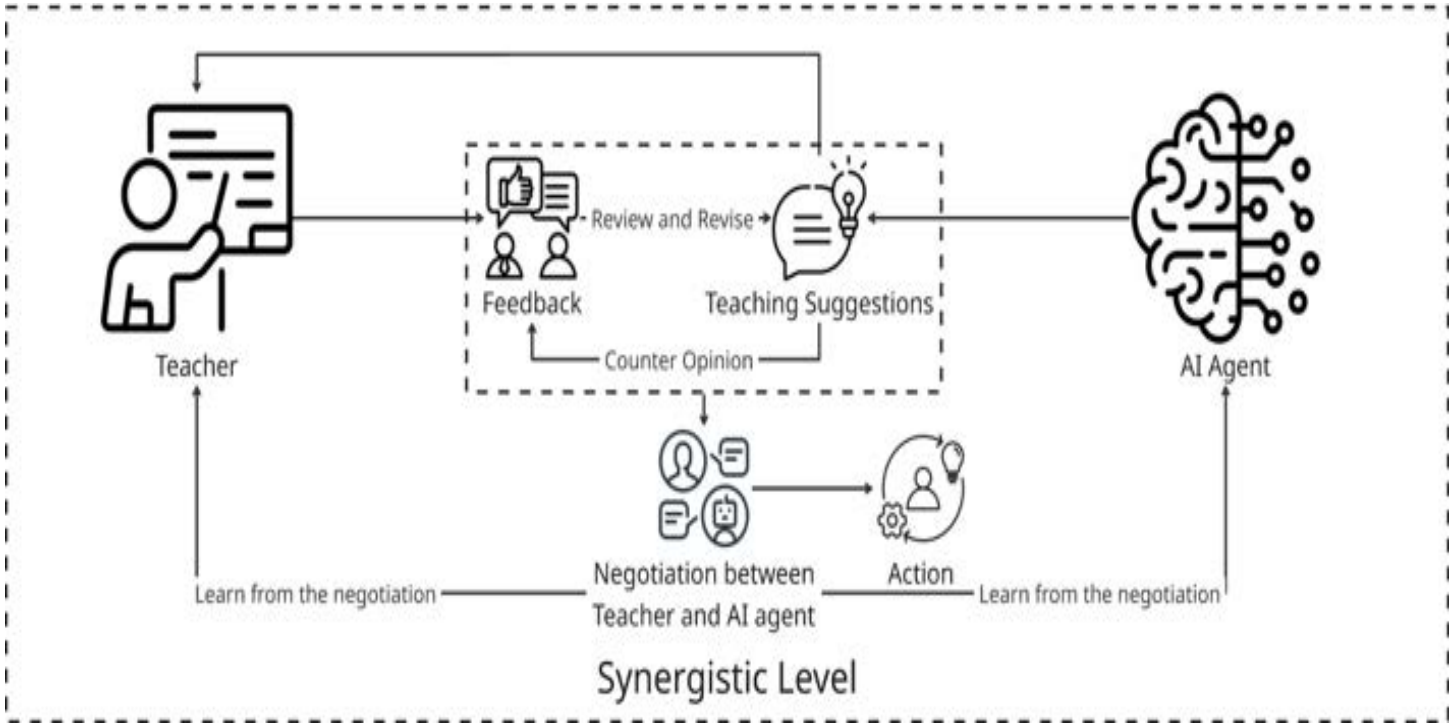
- On average, human–AI teaming significantly better than the human alone (**complementarity**); yet performs **significantly worse** than the best of humans or AI alone **(augmentation).**

- **Task dependent:** Decision making vs content creation, for instance.

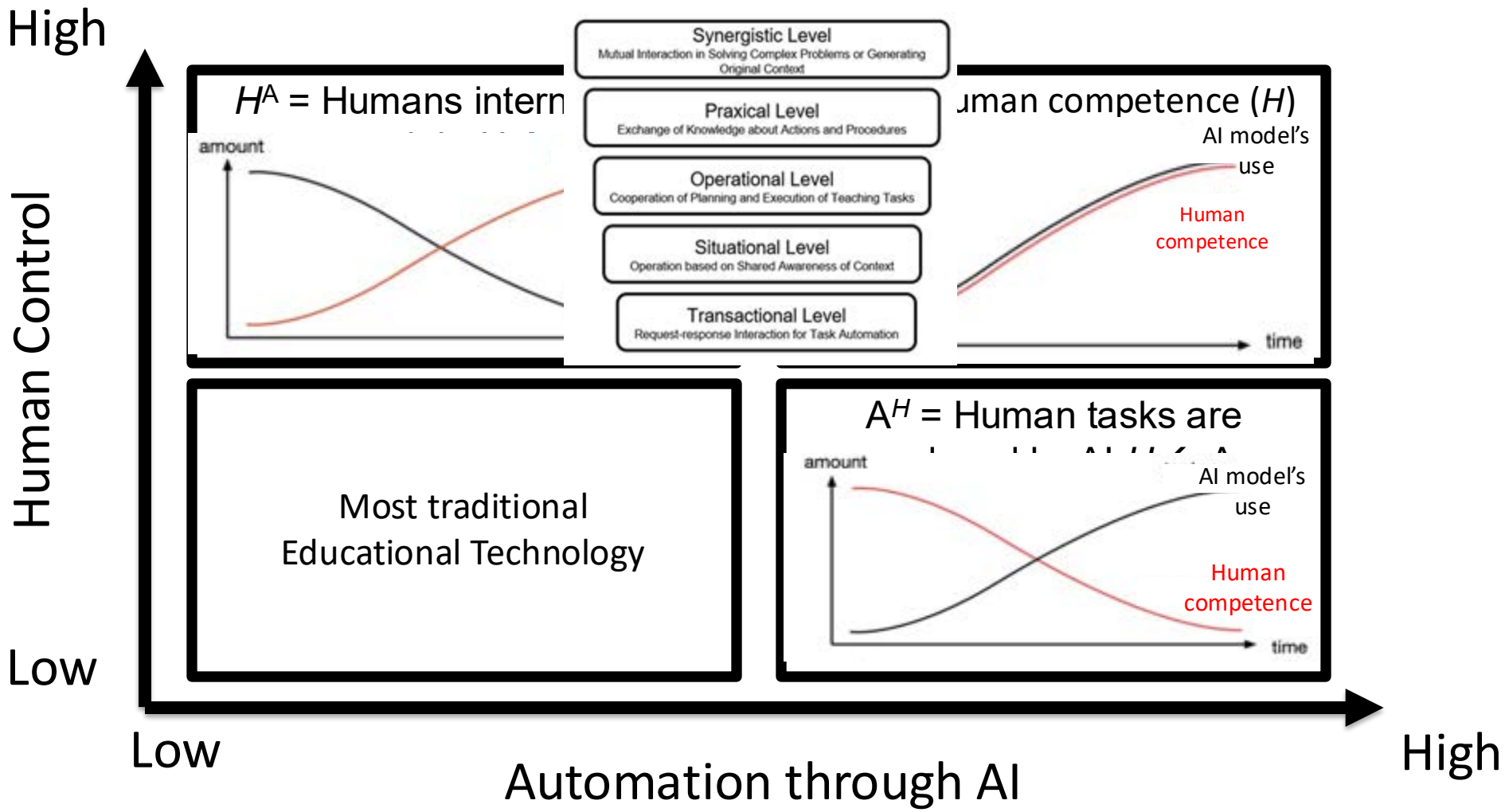


Vaccaro, M., Almaatouq, A., & Malone, T. (2024) When combinations of humans and AI are useful: A systematic review and meta-analysis, *Nature Human Behaviour,* https://doi.org/10.1038/s41562-024-02024-1

From Praxical to Synergistic



AI in Education: A vision for the future



Cukurova, M. (2024). The Interplay of Learning, Analytics, and Artificial Intelligence in Education. *British Journal of Educational Technology*. https://doi.org/10.48550/arXiv.2403.16081

Thank you

Professor Mutlu Cukurova
University College London
m.cukurova@ucl.ac.uk

Taico
Teacher-AI Complementarity

Knowledge Lab UCL UCLAIT - Team