

# Learning and Regulating with ChatGPT: What Experimental Study Tells Us?

+  
Yizhou Fan

Peking University

This study is funded by 2023 SoLAR ECR Grant and Peking University, project: Measuring and Scaffolding Hybrid Human-AI Regulation: Comparing Learning Processes Facilitated by ChatGPT and Human Experts

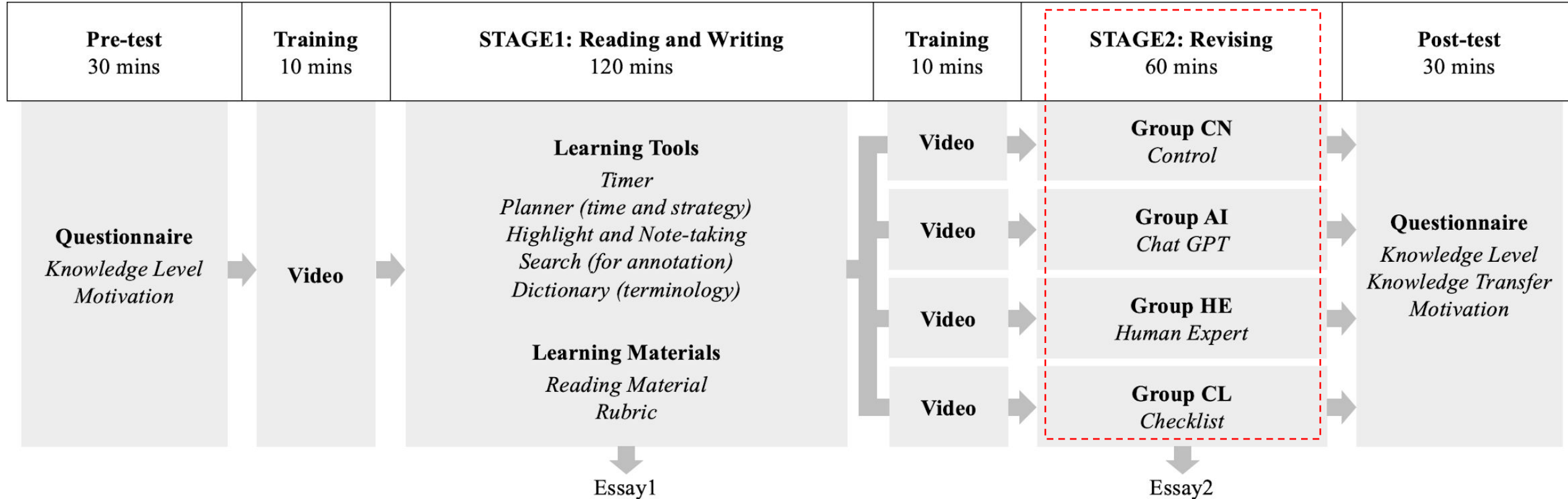
# Early Career Research Grant



- + Enhancing Learning with Learning Analytics in an AI Powered World
- + March 2023 -> lots of challenges, potentials and unknowns about GAI
- + FLoRA project -> Self-regulation to Hybrid Human-AI Regulation
- + Exploratory, experimental, comparative research about ChatGPT
  
- + Note: we are comparing AI and human, but that's not the main aim;
- + The aim is to better understand the relationships of AI and human.

# Experimental design

- 133 participants (university students), AI in education, writing task



Learning content

Navigation menu

3: Future Learning\_1 / General Instructions

**General Instructions**

In this learning session, the goal is to write a vision essay that describes the future of education. Please describe, in 200 to 400 words, how you envision learning in a school in 2035.

Please consult the materials in this learning environment that provide information about three important topics for envisioning the future:

1. Artificial intelligence and its applications
2. What differentiation is and how it affects learning
3. The process of scaffolding a learning experience

The goal of the learning session is to write a vision essay about the future of education in a school in 2035.

At the end of the learning session, you will be asked to:

- explain the concepts of artificial intelligence and its applications
- explain how they affect learning and education
- apply them in the context of a school in 2035
- combine the concepts into a vision of the future of education

For more information about the topics, please refer to the materials in this learning environment.

- Notes Panel
- Search Tool
- Essay tool
- Planner Tool
- Dictionary Tool
- Timer
- Checklist Tool
- ChatGPT Tool
- Ask Teacher Tool

**Essay Writing** 0 words

Write essay here...

Save Essay

(Stage 1)

CN Revise essay by themselves



AI



HE

**Checklist Tool**

Basic Academic Originality Integration and elaboration

Sentence Text: Artificial intelligence has been rapidly developing for years, leading to the emergence of a wide variety of products.  
 Error Position: variety  
 Error Correction: variety  
 Error Explanation: The correct spelling of the word is "variety."  
 Error Type: Spelling

Sentence Text: However, limited extent in tangible AI-based educational products can be seen.  
 Error Position: limited extent  
 Error Correction: a limited extent  
 Error Explanation: The phrase should include an article "a" to make it grammatically correct.

This analysis may take around 5-10 min to process. **Analyse**

**ChatGPT Tool**

could you give me some advice on the academic writing skills of my essay?  
16:12:08

Certainly! Here are some suggestions to improve the academic writing skills in your essay:

1. Use clear and concise language: Aim for clarity in your writing by choosing precise words and avoiding unnecessary jargon or overly complex language. Make sure each sentence contributes to your overall argument.

Ask a question... **Send**

**Ask Teacher Tool** Connected to server: Teacher is typing...  
Teacher online status: Online

Hello  
16:42:38

hi  
16:42:41

could you give me some advice on the structure of my essay?  
16:43:04

Ask a question... **Send**

CL

**Checklist Tool**

Basic Academic Originality Integration and elaboration

Essay Sentence: Nowadays, the role of AI in education is left in some tangible educational products.  
Integration Level: Remember and understanding

Essay Sentence: More and more data are needed to make AI work better.  
Integration Level: Evaluate and create

Essay Sentence: And the scaffolding is that the work is shared alone between the learner and some more knowledgeable other or agent.  
Integration Level: Evaluate and create

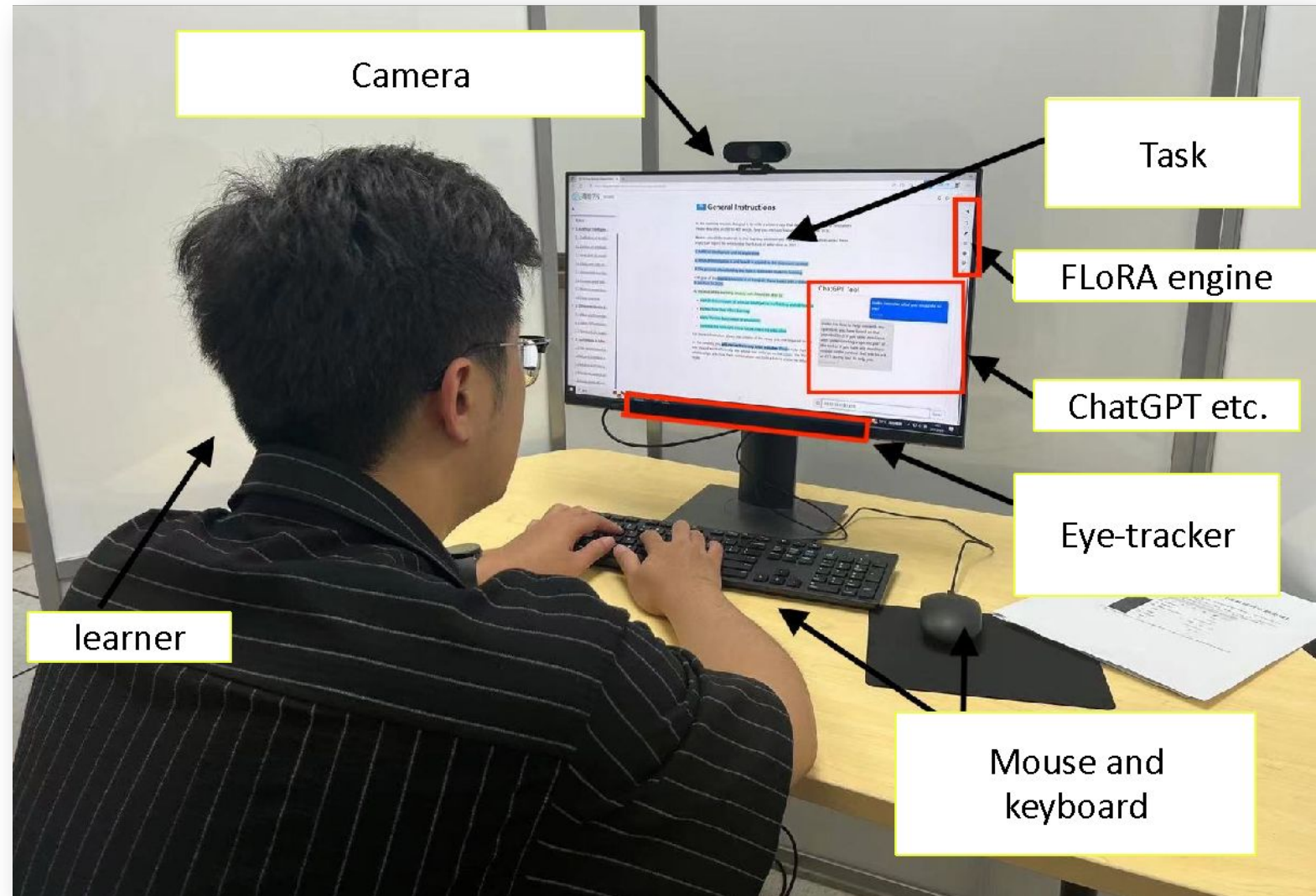
Essay Sentence: In addition, it also enables learning from the experience.  
Integration Level: Apply and analyse

Essay Sentence: Scaffolding to optimize learning needs to acquire more knowledge.  
Integration Level: Apply and analyse

This analysis may take around 5-10 min to process. **Analyse**

# Data collection

- Pre-Post test
- Pre-Post survey
- Learning trace data
- Eye-tracking and video
- Screen recording
- Dialogical text
- Post-study interview
- Etc.



# Learning performance: essay scores

- Essay version 1 (after stage 1), essay version 2 (after stage 2)
- Essay scores improvement (essay version 2 ~ 1)

Comparison	Mean Difference	Lower Bound(95% CI)	Upper Bound(95% CI)	p-adjusted
cl-ai	-2.200	-4.033	-0.367	0.012
cn-ai	-1.970	-3.858	-0.083	0.037
he-ai	-2.120	-4.049	-0.191	0.025
cn-cl	0.230	-1.725	2.184	0.990
he-cl	0.080	-1.915	2.075	1.000
he-cn	-0.150	-2.195	1.895	0.998

- AI group outperforms CN, CL and HE groups

# Learning performance: knowledge gain and transfer

- Knowledge gain (pre-post-test): no significant differences
  - The ANOVA results show no significant differences between four groups in both the pre-test score ( $F=1.294$ ,  $p=0.281$ ,  $\eta^2=0.036$ ) and post-test score ( $F=0.913$ ,  $p=0.438$ ,  $\eta^2=0.030$ ).
- Transfer test (AI in medical science): no significant differences
  - ANOVA results show no significant differences between four groups ( $F=0.019$ ,  $p=0.996$ ,  $\eta^2=0.000$ ).

# Intrinsic motivation (IMI, McAuley et al. 1989)

- No significant difference between the four groups was observed:
  - Interest/Enjoyment ( $F=1.087$ ,  $p=0.358$ ,  $\eta^2=0.029$ );
  - Perceived Competence ( $F=0.453$ ,  $p=0.716$ ,  $\eta^2=0.012$ );
  - Effort/Importance ( $F=1.152$ ,  $p=0.332$ ,  $\eta^2=0.030$ ) and
  - Pressure/Tension ( $F=0.546$ ,  $p=0.652$ ,  $\eta^2=0.015$ ).
- Although the insignificant were observed, we found:
  - CN group reported lowest interest and enjoyment, and highest pressure and tension
  - CL group reported highest scores for interest and enjoyment, perceived competence and effort, while they reported lowest pressure and tension

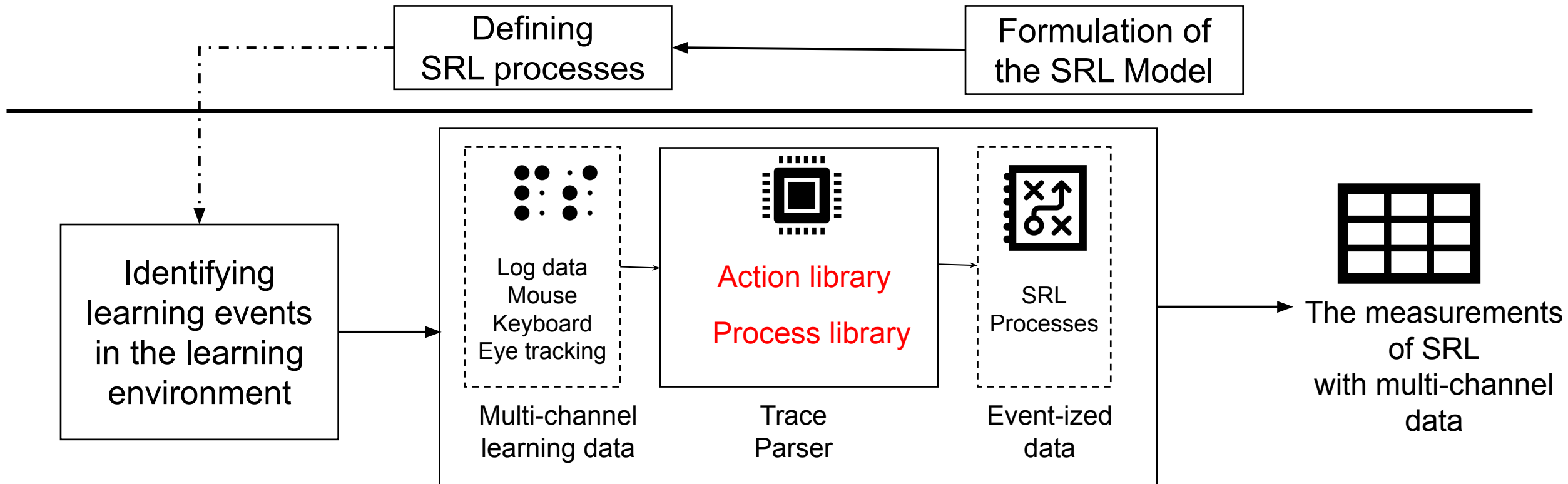


# SRL processes

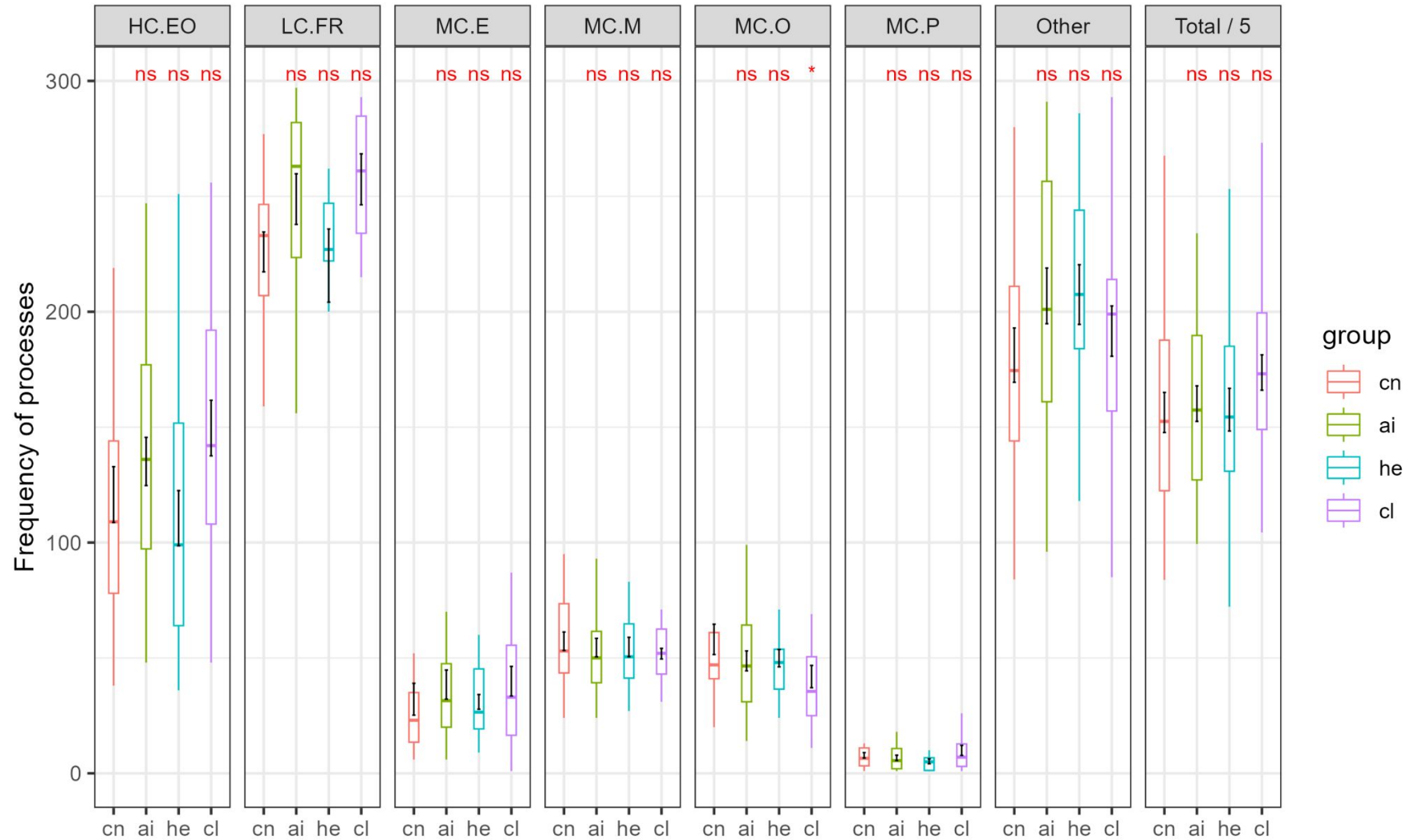
## *FLoRA* trace parser

<https://floraproject.org/website/>

Main Categories	Subcategories	Codes
Metacognition	Orientation	MC.O
	Planning	MC.P
	Monitoring	MC.M
	Evaluation	MC.E
Low_Cognition	First-reading	LC.F
	Re-reading	LC.R
High_Cognition	Elaboration/Organisation	HC.E/O

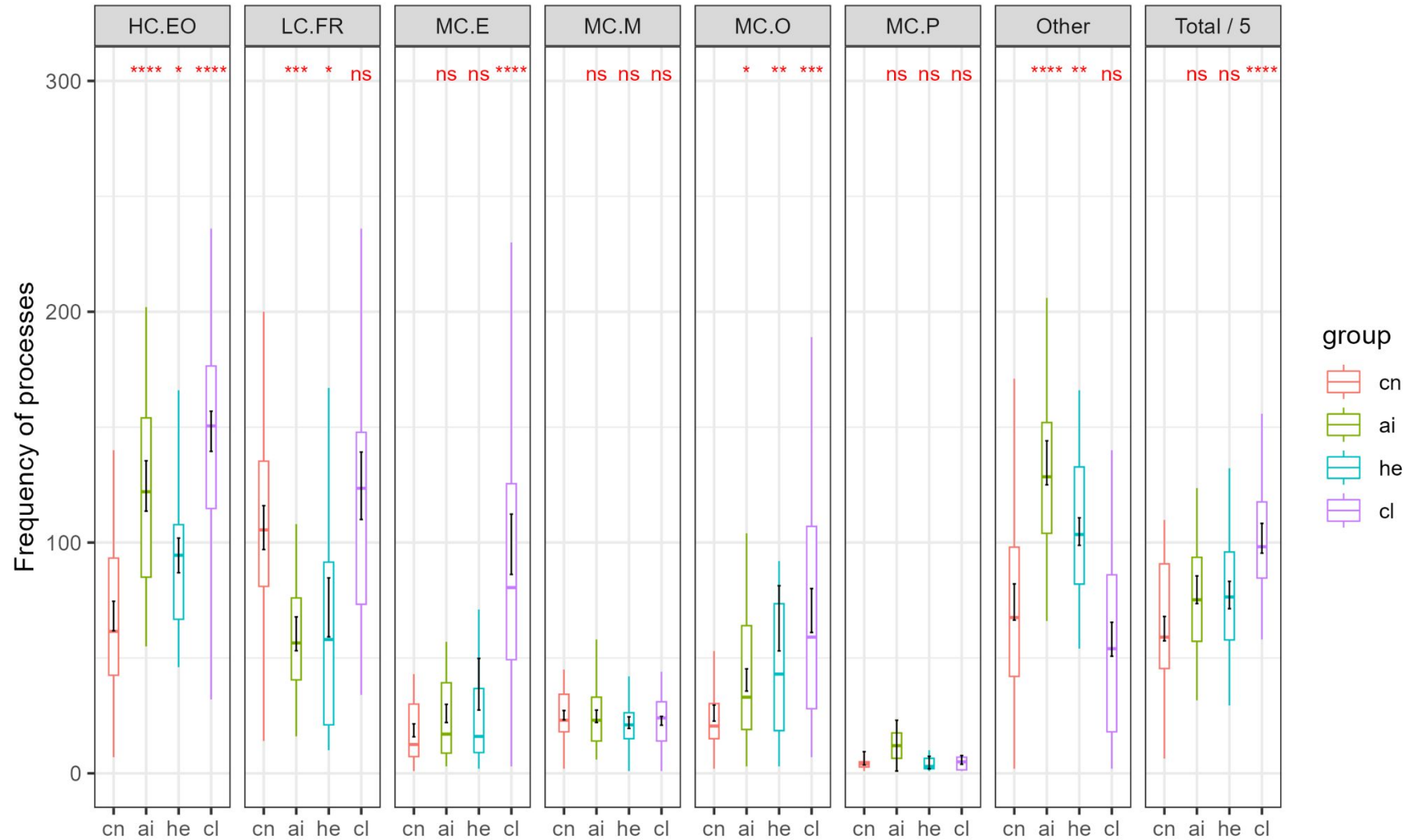


# Stage 1



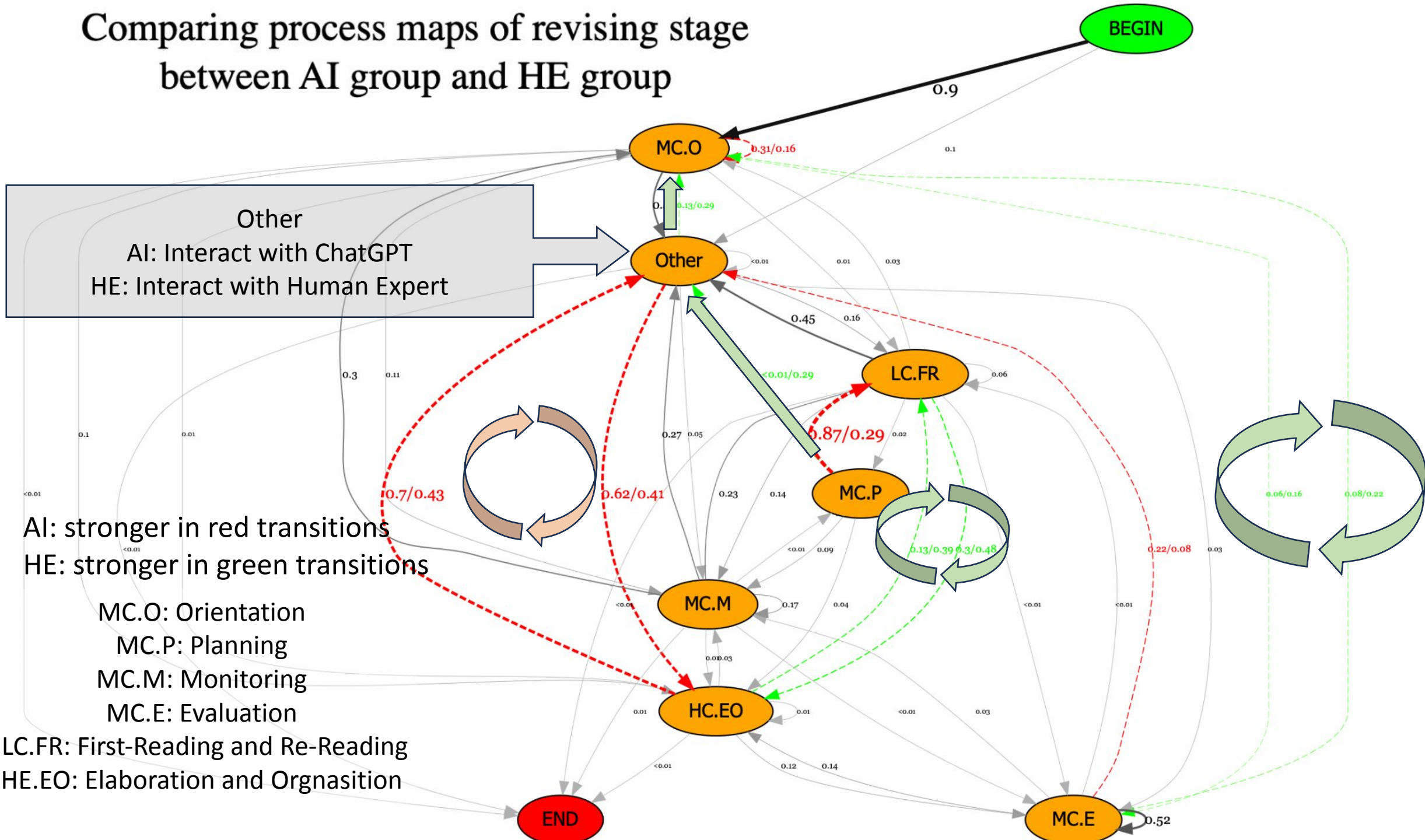
Comparing learning processes of the first learning stage (2 hours reading and writing) among four groups

# Stage 2



Comparing learning processes of the second learning stage (1 hour revising) among four groups

# Comparing process maps of revising stage between AI group and HE group



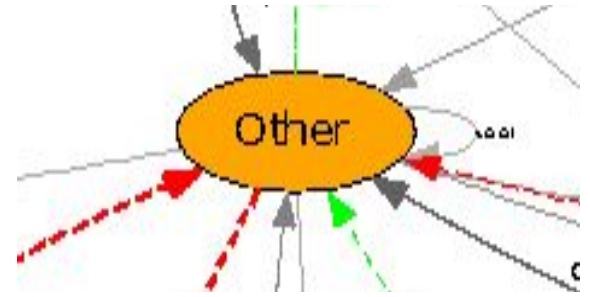
# What Experimental Study Tells Us? 1st Insight

- + High-intelligence tools (such as ChatGPT) may not stimulate intrinsic motivation to learn and knowledge gain/transfer, but can rapidly improve short-term performance;
- + Potential Metacognition Laziness and over-reliance, and the offloading of (meta)cognitive load can be the two sides of a coin;
- + "AI-empowered learning skills" which optimises performance at the expense of developing genuine human skills (?)
- + One important note: ChatGPT is excellent at utilizing clear rubrics

Fan et al., Beware of Metacognitive Laziness: Effects of Generative Artificial Intelligence on Learning Motivation, Processes, and Performance, British Journal of Educational Technology (under review)

Stage	Activities	Definition	Code
-------	------------	------------	------

Try to open this black box?



Diagnosing Question	Diagnosing Question	learners need to determine if there is a problem based on their own learning status and decide if they need help	Diag.Ques
---------------------	---------------------	--	-----------

Asking Help	Asking Help.Instrumental	Instrumental help-seeking (learners ask facilitated hints that can assist them in revising their essays independently afterward)	Ask.Instr
-------------	--------------------------	--	-----------

Asking Help.Executive	Executive help-seeking (learners tend to look for answers that can be applied directly)	Ask.Exec
-----------------------	---	----------

Code the screen recording

Evaluating Help	Asking Help.Avoidant	Avoid asking for help (learners try to ask questions but not to send them)	Ask.Avo
	Evaluating Help.Positive	Positive evaluation (learners evaluate the help messages and give positive feedback)	Eva.Pos

Evaluating Help.Negative	Negative evaluation (learners evaluate the help messages and give negative feedback or not to give any feedback)	Eva.Neg
--------------------------	--	---------

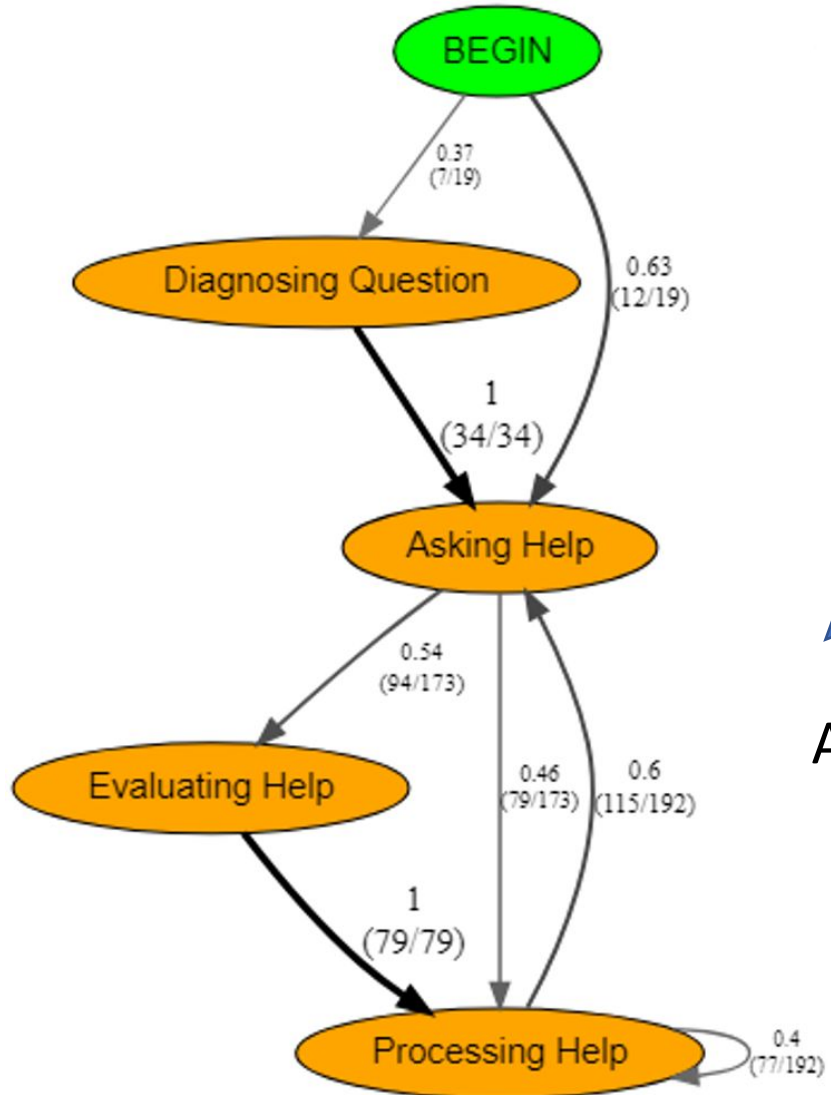
Help-seeking process model (Nelson-Le Gall, 1981)

Processing Help	Processing Help.Accepting	Accepting help (learners apply the help directly in revising their essays)	Pro.Acc
-----------------	---------------------------	--	---------

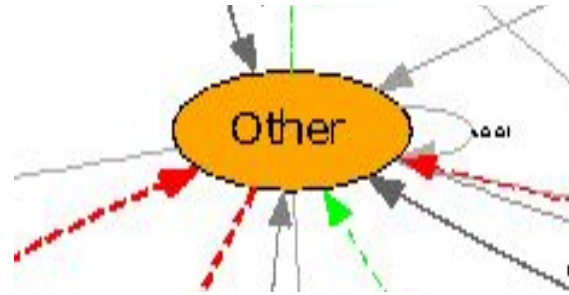
Processing Help.Neglecting	Neglecting help (learners do not apply the help directly in revising their essays)	Pro.Neg
----------------------------	--	---------

Processing Help.Returning	Returning to the help (learners look back to some previous help message)	Pro.Re
---------------------------	--	--------

# Help-seeking

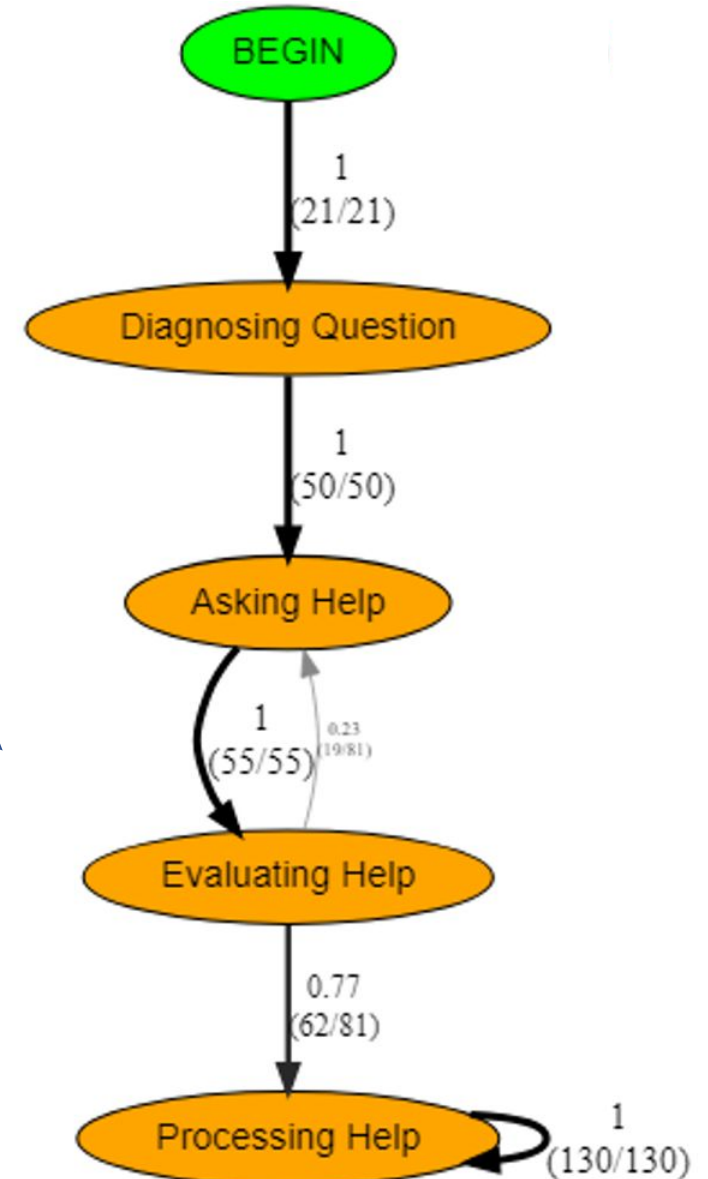


Try to open this black box?



Ask ChatGPT

Ask Teacher



# Comparison of activities between AI and HE Group

Comparison of activities between AI Group and HE Group by Mann-Whitney *U*

Activities	Mean Ratio in cognitive and behavioral activities (%) (AI, N=18)	Mean Ratio in cognitive and behavioral activities (%) (HE, N=20)	Mean Rank (AI, N=18)	Mean Rank (HE, N=20)	Z	Effect Size (ES)	Sig. (2-tailed)
Ask.Instr	64.75	77.86	22.83	16.50	-1.552	-0.183	.121
Ask.Exec	31.93	6.05	25.28	14.30	-3.141	0.756	.002**
Ask.Avo	3.30	16.08	17.00	21.75	-1.657	0.128	.098
Eva.Pos	2.22	54.55	10.83	27.30	-4.937	-0.778	.000***
Eva.Neg	97.77	45.44	24.22	15.25	-4.937	0.778	.000***
Pro.Acc	64.92	60.41	19.58	19.43	-1.142	0.256	.253
Pro.Neg	8.42	3.24	23.53	15.88	-2.231	0.678	.026*
Pro.Re	26.65	36.33	15.36	23.23	-2.373	-0.417	.018*



- Instructions
- General Instructions
- Rubric
- 1: Artificial Intelligen...
  - 1.1 Definition of Artifici...
  - 1.2 History of Artificial ...
  - 1.3 How does AI work?
  - 1.4 Ethics and risks of...
  - 1.5 Supervised machin...
  - 1.6 Unsupervised mach...
  - 1.7 Reinforcement lear...
  - 1.8 Deep Learning
- 2: Differentiation in E...
  - 2.1 What is Differentiat...
  - 2.2 Using differentiatio...
  - 2.3 Standards for teach...
- 3: Scaffolding in Educ...

### 5: Essay Revision / General Instructions

## General Instructions

In this learning session, the goal is to write a vision essay that describes the future of education. Please describe, in 200 to 400 words, how you envision learning in a school in 2035.

Please consult the materials in this learning environment that provide information about three important topics for envisioning the future of education in 2035.

- Artificial intelligence and its application
- What differentiation is and how it is applied in the classroom context
- The process of scaffolding and how it optimizes students learning

The goal of the learning session is to integrate these topics into a vision of learning in a school in 2035.

At the end of the learning session, you should be able to:

- explain the concepts of artificial intelligence, scaffolding and differentiation
- explain how they affect learning
- apply them in the context of education
- combine the concepts into a future vision for education

For more information about the criteria of the essay, you are required to read the rubric.

In this session, you **will revise the essay in 60 minutes**. Please note that you should work efficiently. We advise you to focus on the [rubric](#).

## ChatGPT Tool

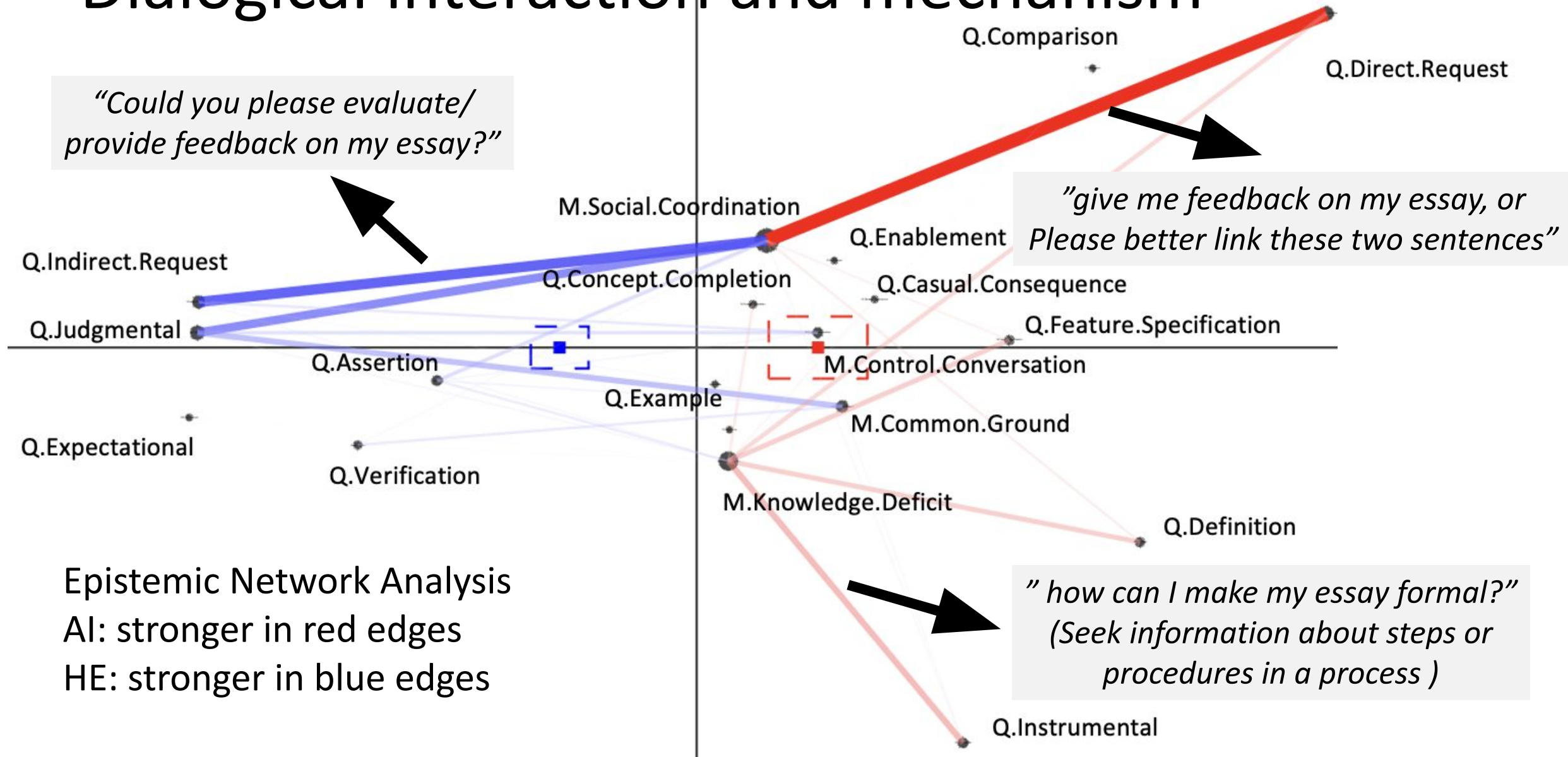
differentiation strategies. This could result in a more inclusive and effective learning environment, where all students are able to reach their full potential.

describe What differentiation is and how it is applied in the classroom context

Processing

Ask a question... Send

# Dialogical interaction and mechanism



# What Experimental Study Tells Us? 2nd Insight

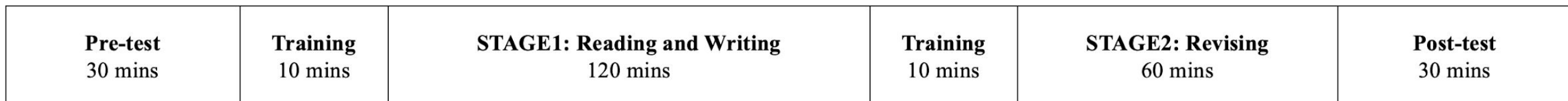
- + Learners ask AI pragmatic questions -> improve their performance;
- + Learners experienced lower social cost in the AI group compared to asking the human expert;
- + Learners showed adaptivity when facing different facilitators;
- + Previous theories and models (e.g., Linear help-seeking process) may encounter difficulties in explaining human-AI interaction;
- + Concerns about learners' lack of evaluation and monitoring when seeking executive help from ChatGPT -> scaffoldings

Chen et al., Unpacking Help-Seeking Processes through Multimodal Learning Analytics: A Comparative Study of Learning Facilitated by ChatGPT and Human Expert (drafting);

Cheng et al., Asking Questions of Generative Artificial Intelligence Improves Academic Performance (drafting)

# But, who did learners prefer to learn with?

- Human-AI preference scale (5 questions, pre and post task)
  - e.g, “Compared to AI, human tutors can better understand the main idea of my article and provide more helpful suggestions.”
- Human-AI choice (pre and post task)
  - 1 -> prefer human over AI; 0 -> prefer AI over human
- Preference Alteration
  - -1 indicates participant altered preference from human tutor to AI;
  - 1 indicates participant has altered preference from AI to human tutor;
  - 0 indicates participant’s preference remains unchanged.



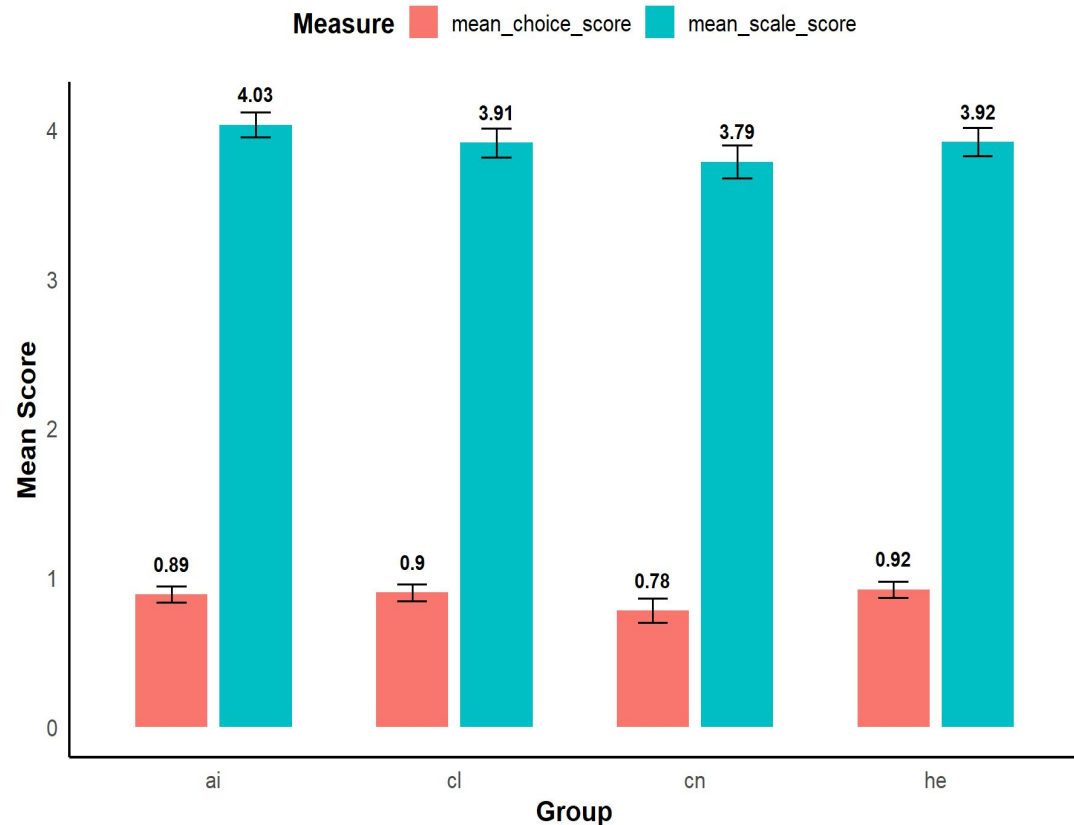
1st preference  
measure

2nd preference  
measure

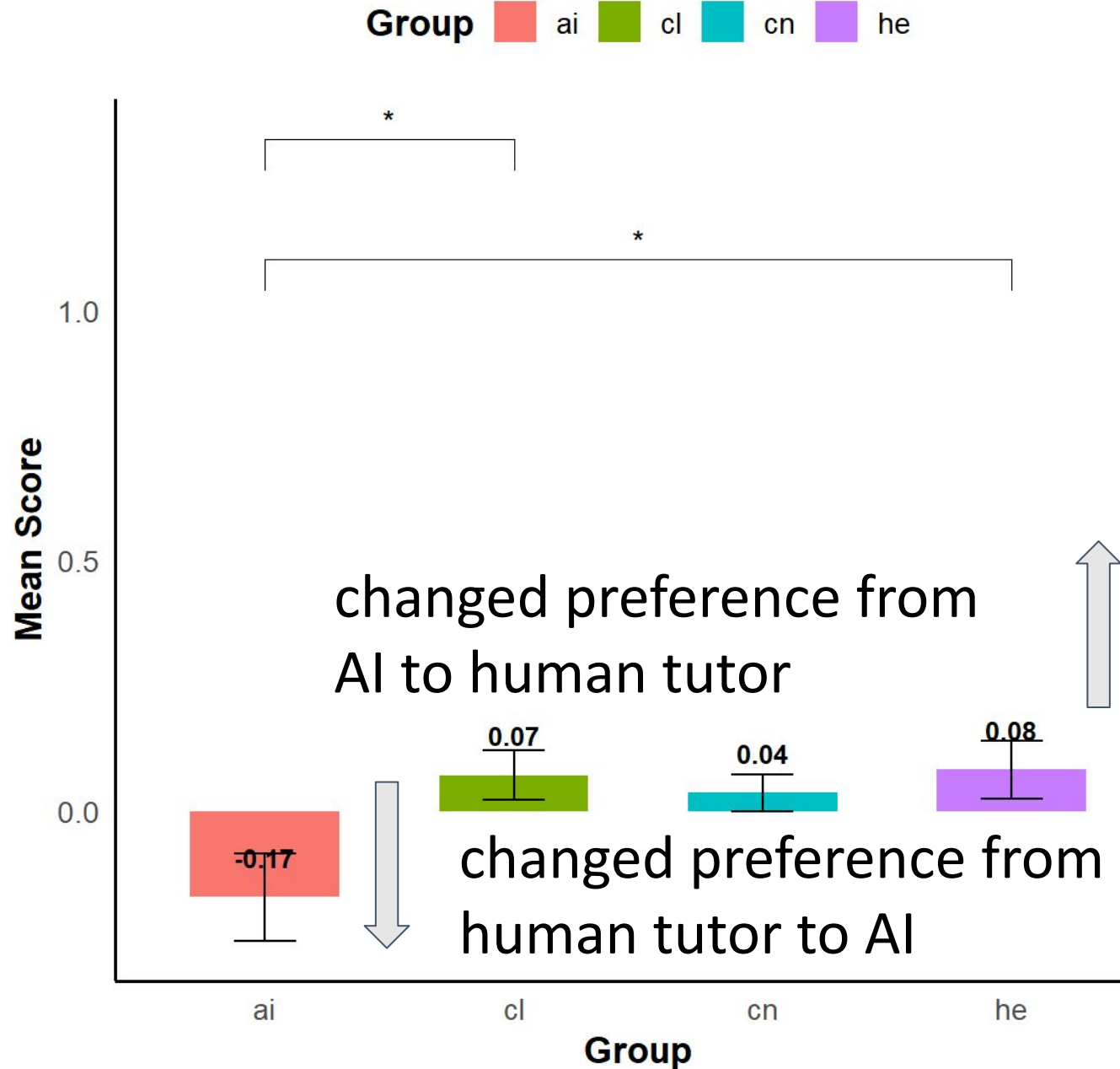
# Human-AI preference

No difference before task

In general, learners prefer human

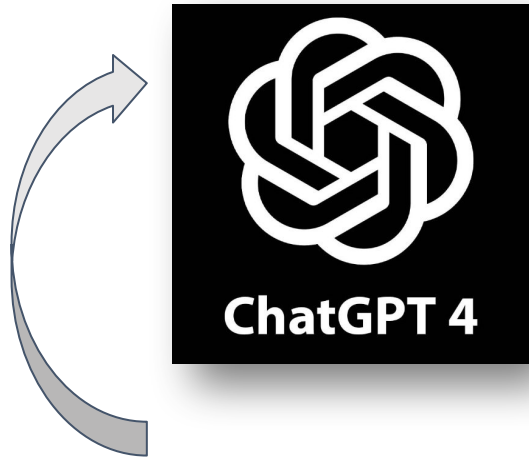


## Preference Alteration Score by Group

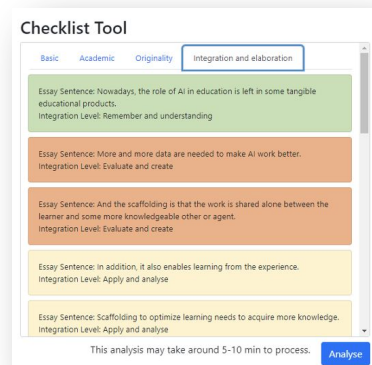


# Polarization of preference change

- AI group who has experienced ChatGPT interaction are more inclined to AI
- HE group who has experienced human expert interaction are more inclined to humans (100% choose humans)
- CN group shows no significant change, very slightly turns to human
- CL group who has not experienced Chatgpt and human expert, surprisingly, also are more inclined to humans (97% choose human)



Rubric



	Principle	Format of feedback
Basic	Check based on GPT4.0	Highlight spelling grammar errors
Academic	Based on a database of academic norms developed by teachers	Highlight Words/phrases with problematic academic style
Originality	Calculation based on similarity of more than 7 words	Highlight repeat sentences
Integration and Elaboration	Division based on Bloom's taxonomy of cognitive domains	Different color highlights represent different levels of processing

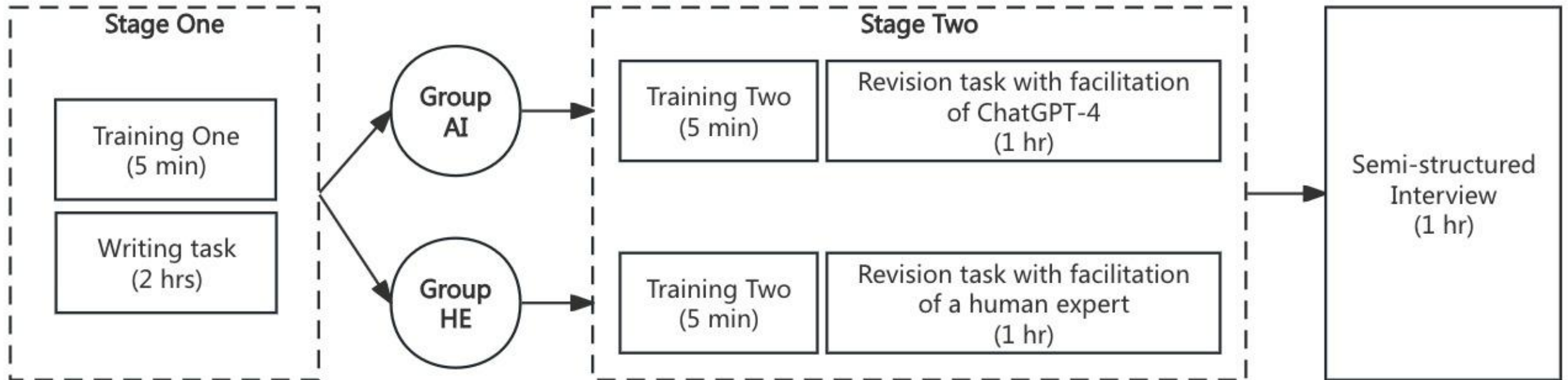
# What Experimental Study Tells Us? 3rd Insight

- + ChatGPT and Checklist: what is the main difference?
- + The dialogic format may be naturally more attractive than traditional feedback tools or dashboard;
- + The intelligence level of AI agents or learning tools affects learners' trust and aversion on algorithms/techniques.
- + No matter how the data show the practical of learner-AI interaction, human (teachers) always have irreplaceable human's value, and learners **STILL** prefer to learn with human expert in our task.

Le et al., Rolling to the edge: investigating learners' preference for learning supports from human-tutor, traditional AI tutor and LLM (drafting)



# Interviews to understand learners' perspective

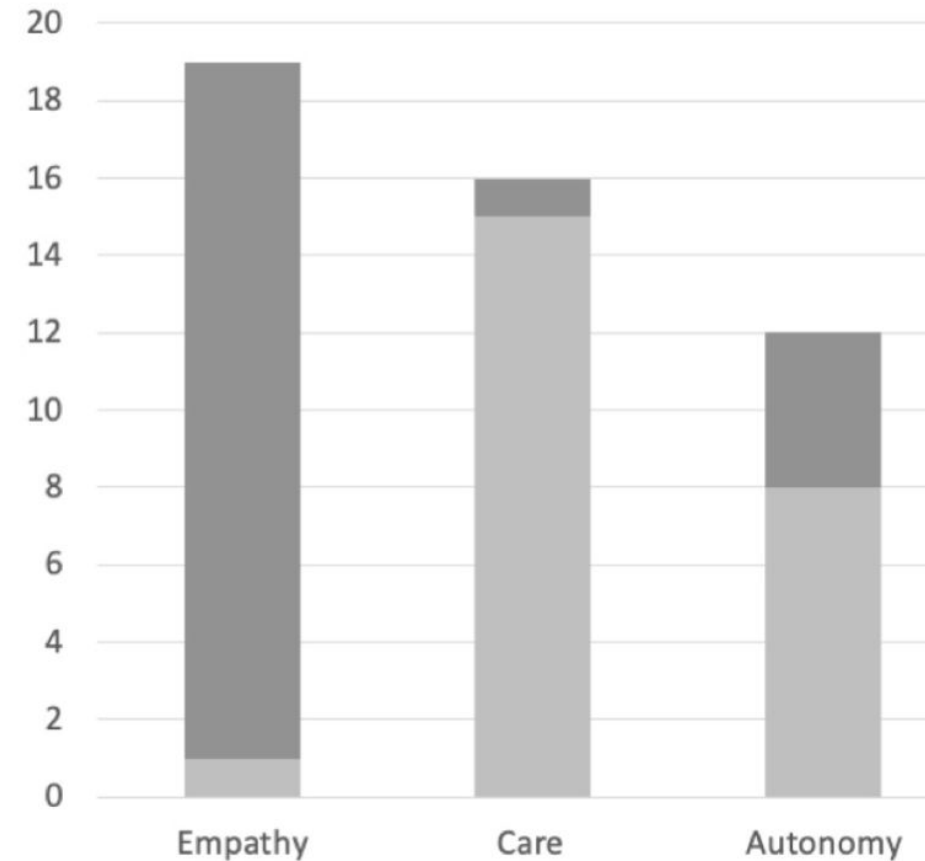


AI group: 33 accepted the interview

HE group: 26 accepted the interview

# Different values and value alignment

Value	Working definitions	Examples
Empathy	comprehend and respond to learner's intentions and emotions properly	AI27: Even if my question was not clear enough, the teacher still knew what I wanted to ask.
Care	alleviate the interpersonal stress of learners and demonstrate patience to them	HE03: I hope the teacher will never get tired and judgmental about my questions.
Autonomy	Respect for learners' freedom to make their own choices and to monitor their learning process	AI13: I have greater power of selectivity and decision-making and greater freedom.



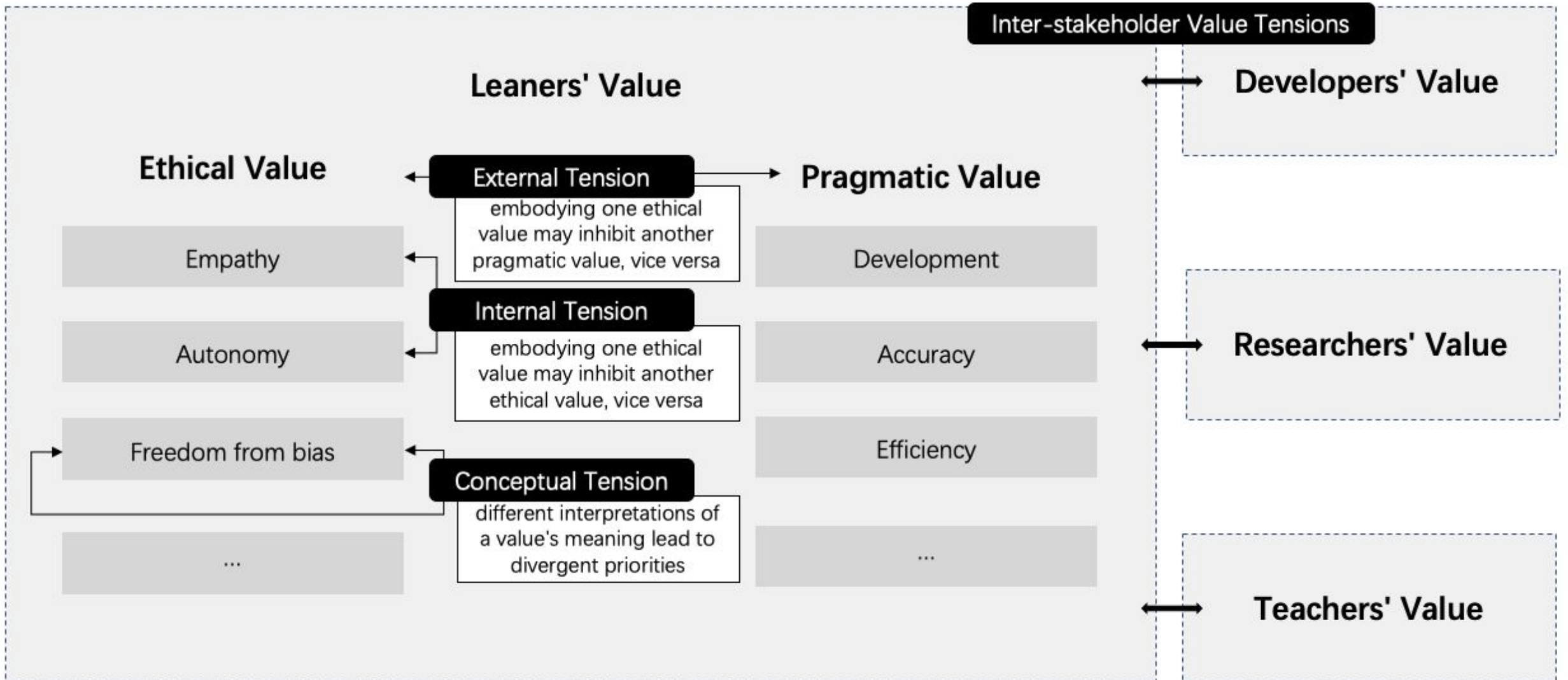
Interview  
Data



■ Better embodied value by teacher (human expert)

■ Better embodied value by AI (ChatGPT)

# Value tensions of learning with AI or human expert



# What Experimental Study Tells Us? 4th Insight

- + Human and AI (e.g., ChatGPT) each have their own unique value;
- + Learners also dynamically perceive and evaluate affordances of different learning facilitators as they regulate their own learning;
- + Value as a key ethic issue of AI in education was relatively neglected
- + Different stakeholders should keep value sensitive design in mind and seek the balance between different values;

Shen et al., Aligning and Comparing Values of ChatGPT and Human as Learning Facilitators: a Value-Sensitive Design Approach, British Journal of Educational Technology (under review)

# Let's get back to the CL group

MC.E			
	ns	ns	****

## Essay Writing

146 words

Normal B I U

With the development of new tech and methods, what will the future of education be like?

Artificial intelligence is the ability of computers to perform tasks that require humans to use their intelligence. It is "an approach to teaching in which teachers proactively modify curricula, teaching methods, resources, learning activities, and student products to address the diverse needs of individual students and small groups of students to maximize the learning opportunity for each student in a classroom". Furthermore, teachers may not only take into account differences in students' cognitive abilities, but also other differences such as in students' motivation or interest for example. In addition, there is a concept called "scaffolding". It is to describe how children, with the help of someone more knowledgeable to share and support their problem solving, can perform more complex tasks than they would otherwise be capable of performing on their own.

Save Essay

## Checklist Tool

Basic Academic Originality Integration and Elaboration

Essay Sentence: With the development of new tech and methods, what will the future of education be like?  
Integration Level: Remember and understanding

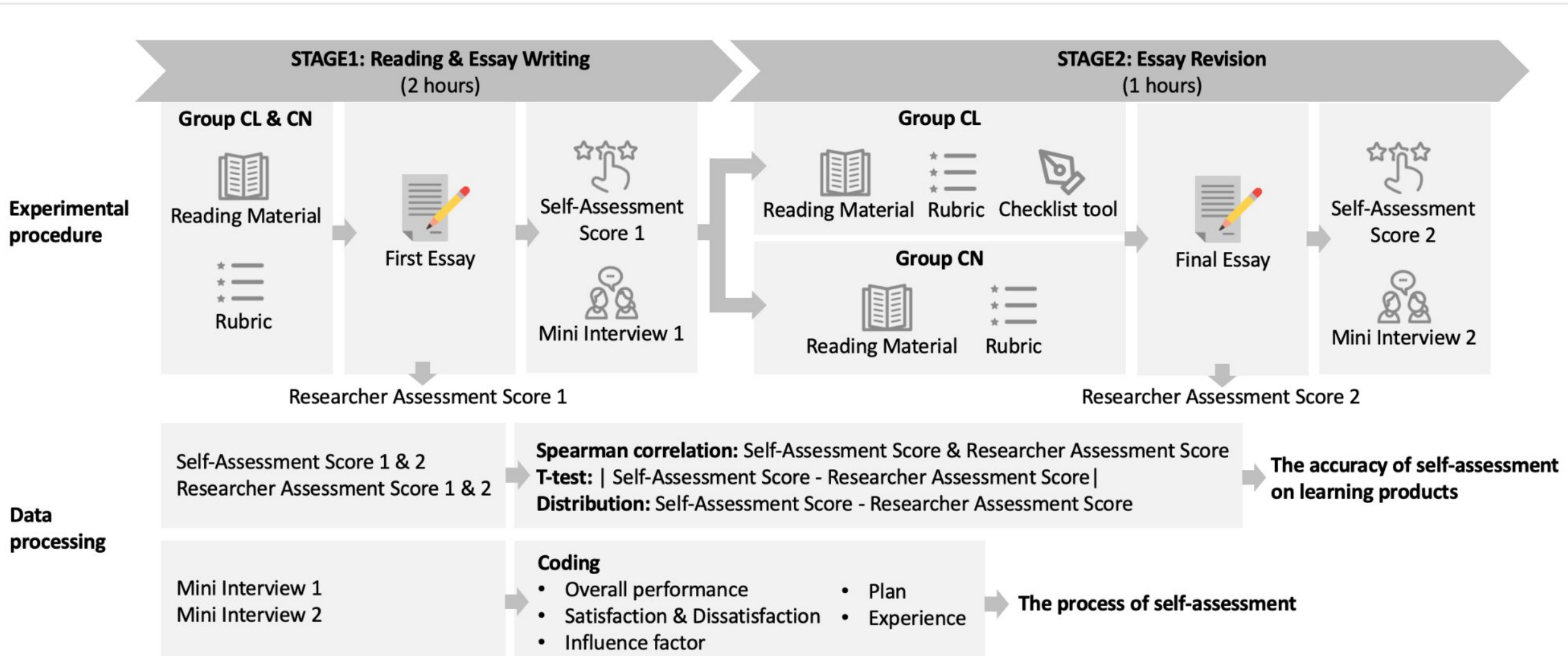
Essay Sentence: Artificial intelligence is the ability of computers to perform tasks that require humans to use their intelligence.  
Integration Level: Evaluate and create

Essay Sentence: It is "an approach to teaching in which teachers proactively modify curricula, teaching methods, resources, learning activities, and student products to address the diverse needs of individual students and small groups of students to maximize the learning opportunity for each student in a classroom".  
Integration Level: Apply and analyse

Analysis

cn ai he cl

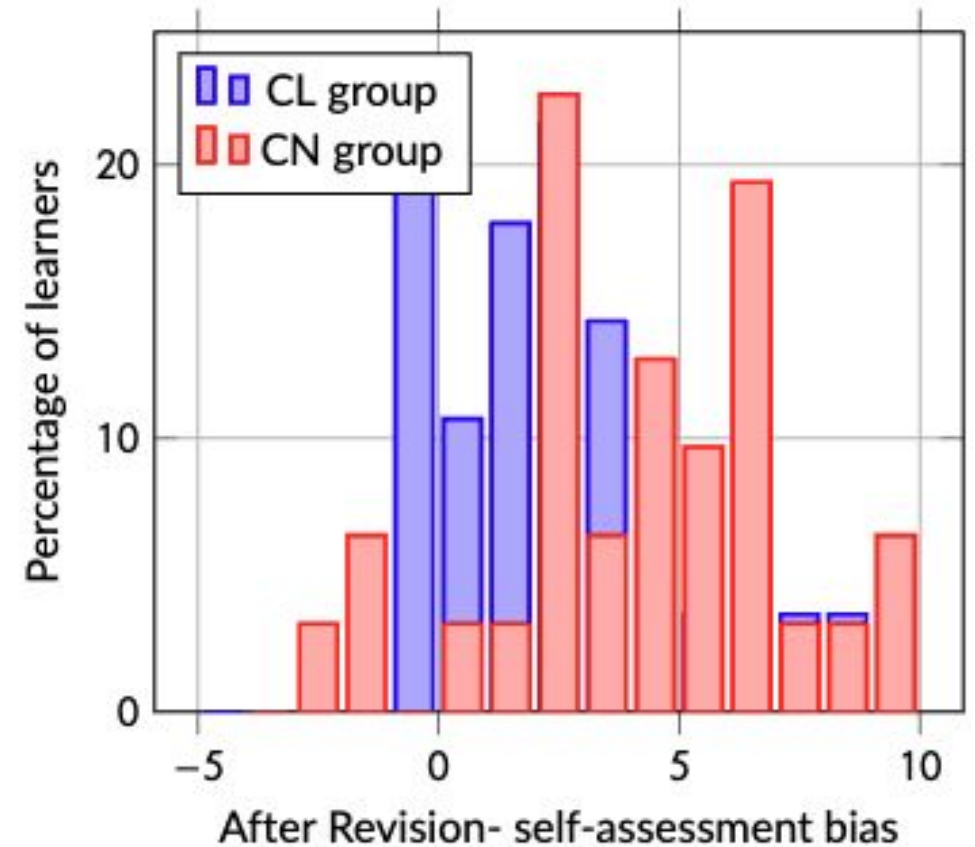
# Trigger and adjust self-assessment



Correlations between SA-Scores and RA-Scores.

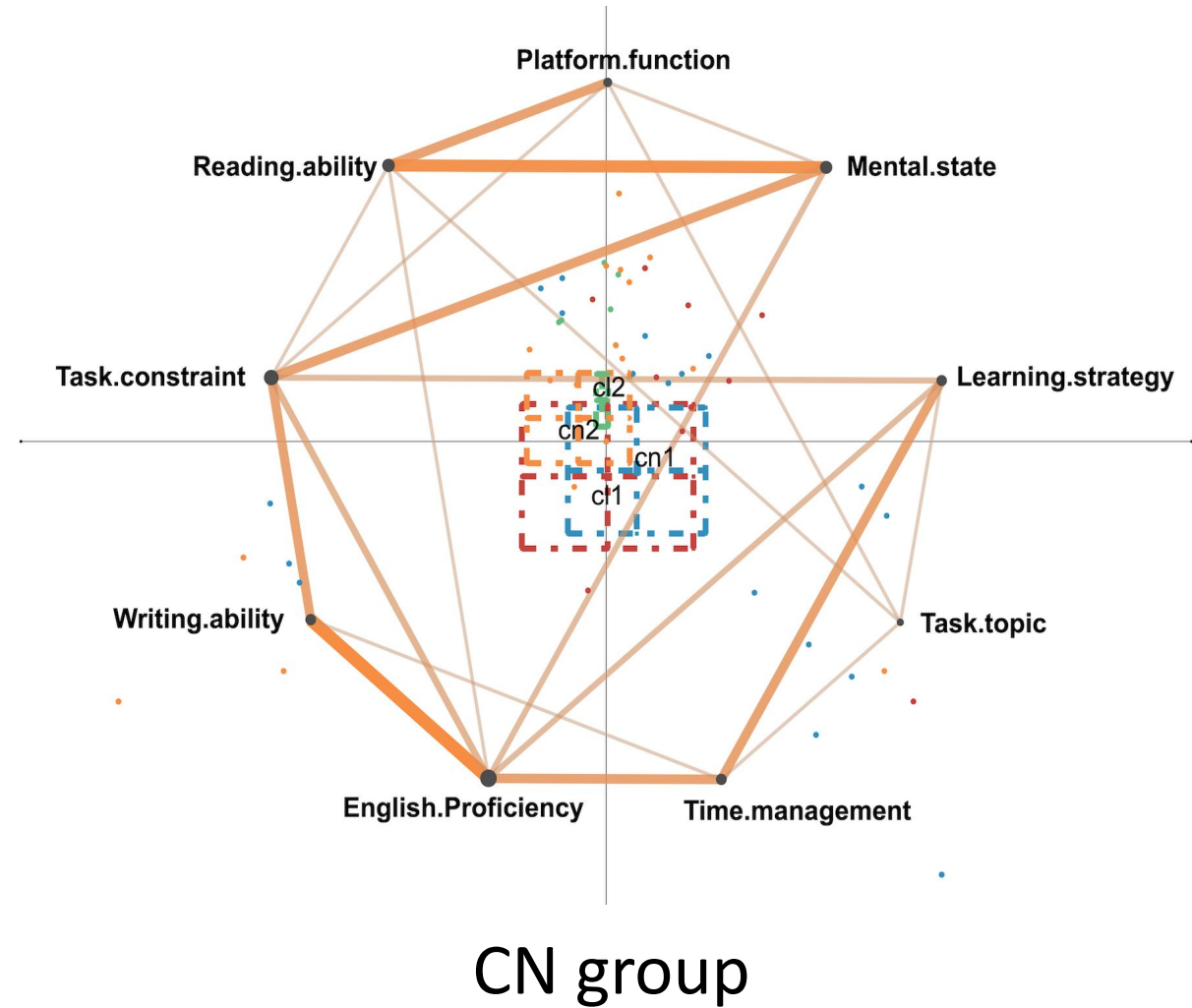
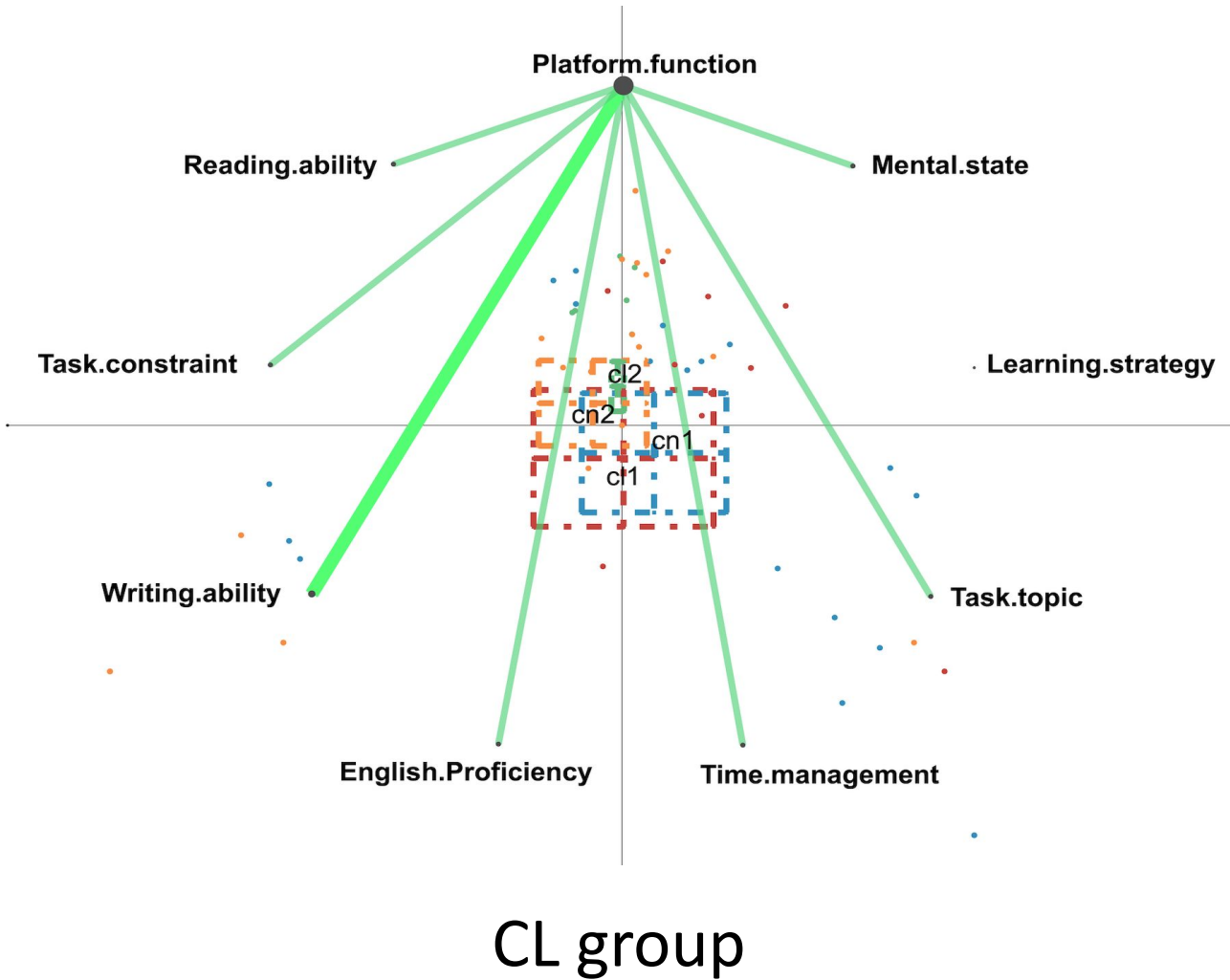
CN group	Before Revision	After Revision
Basic Writing skills	0.384 <sup>a</sup>	0.162
Academic Writing skills	0.155	0.403
Originality	0.179	-0.202
Integration of three topics	0.087	-0.158
Future vision on education	0.206	-0.136
Total score	0.071	-0.453 <sup>b</sup>
CL group	Before Revision	After Revision
Basic Writing skills	-0.037	0.302
Academic Writing skills	0.009	0.229
Originality	0.246	0.142
Integration of three topics	0.495 <sup>b</sup>	0.132
Future vision on education	0.206	0.101
Total score	0.213	0.431 <sup>b</sup>

*a* :  $p < 0.05$ , *b* :  $p < 0.01$



Checklist tools improved self-assessment, and performed better than CN (and AI and HE) groups

# Factors that influenced writing: learners' feedback





# What Experimental Study Tells Us? 5th Insight

- + Writing analytics feedback tools or dashboards also have unique values;
- + Adjusting learners' self-assessment is a core part of self-regulated learning and has been under-researched in previous studies;
- + Low affordability tools (such as bloom taxonomy tool in the Checklist toolkit) are not very useful for self-regulated learning;
- + When tools provided, regardless of their affordance, learners will subjectively consider tools as the primary factor, thereby inhibiting reflection on their own abilities.

# Discussion: Human and AI (人和人工智能)

At the 2008 Beijing Olympics Opening Ceremony, thousands of actors spent five minutes performing only one Chinese character: 和 (hé)



Conjunction or preposition: and, with

Noun: sum, peace, etc

Verb: mix, agree, join, blend, fellow, etc

Adjective: gentle, moderate, harmonious

# Discussion: the rich connotations of “和”

- + Learning and regulating with ChatGPT, but NOT simply using AI as a tutor to replace human teacher;
- + Learning followed, joined and moderated by AI;
- + Future learning and teaching of combining human and AI;
- + Hybrid intelligence: learner-AI, teacher-AI, learner-teacher-AI, etc;
- + The interplay of human control and AI automation (Cukurova, 2024);
- + Scaffolding high human and AI-empowered skills (Gasevic, 2024);
- + Different stakeholders learn, work and live in harmony with artificial intelligence!

# Future works: AI-scaffolded dialogue space

The screenshot displays a digital learning environment for a medical simulation. On the left, a sidebar lists course content, with 'Artificial Intelligence SP C...' selected. The main area shows a 'WEB PAGE' titled 'Artificial Intelligence SP Consultation' with a 'Web page' tab. The case scenario includes a scene from 2023-06-14 at an emergency department, a role for an internal medicine emergency doctor, and patient information. An 'Essay writing tool' window is open, showing a word count of 0 and a text area with the following content:  
Possible diagnoses: Angina, Pulmonary Embolism, Pericarditis, Aortic Dissection  
Preliminary diagnosis: Aortic Dissection  
Differential diagnoses: Myocardial Infarction, Pulmonary Embolism  
A 'Save composition' button is located at the bottom of the tool. On the right, a chat window titled 'patient' shows a dialogue between a doctor (Yan) and a patient. The patient's messages are in blue bubbles, and the doctor's responses are in grey bubbles. The chat history includes:  
Patient: 'Is it just your chest? Do you have radiating pain?' (12:24:05)  
Doctor: 'My back hurts too, and so does my waist.' (12:24:08 Round 20)  
Patient: 'Is your pain like a tearing sensation?' (12:24:17)  
Doctor: 'It feels like I'm being torn apart.' (12:24:19 Round 21)  
Patient: 'Where did the pain start?' (12:24:21)  
Doctor: 'It started in my chest.' (12:24:23 Round 22)  
Patient: 'After the chest pain, I quickly felt the pain in my back as well.' (12:24:25)  
The chat interface includes a search bar, a 'Send' button, and a 'Ask a question...' input field.

# I'm hiring!

Please consider join our team if you are looking for post-doc position and find our studies interesting! email -> [fyz@pku.edu.cn](mailto:fyz@pku.edu.cn)



[ABOUT](#)

[MEMBERSHIP](#)

[EVENTS](#)

[NEWS](#)

[PUBLICATIONS](#)

[COMMUNITY](#)

## Solar Job Board

The go-to place to find positions related to Learning Analytics.

# Thank you!

Yizhou Fan

fyz@pku.edu.cn

Peking University