Collect classroom videos from various platforms (2 languages and 3 subjects)

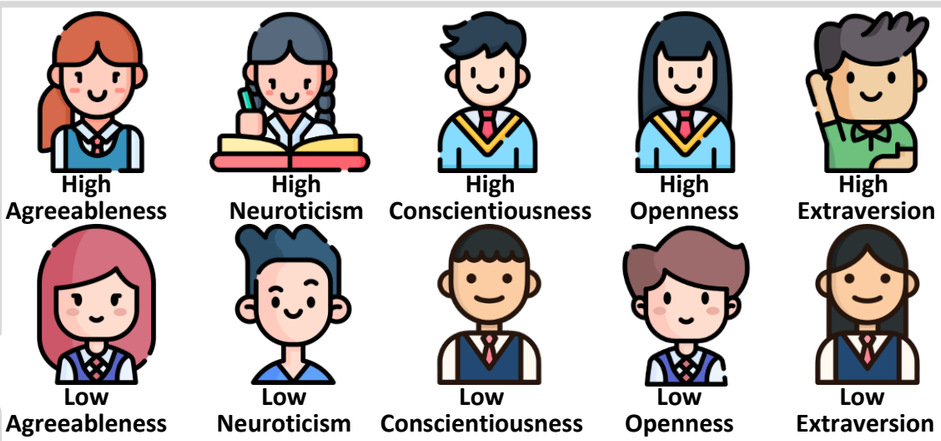
Chinese
401 dialogue turns

Mathematics
395 dialogue turns

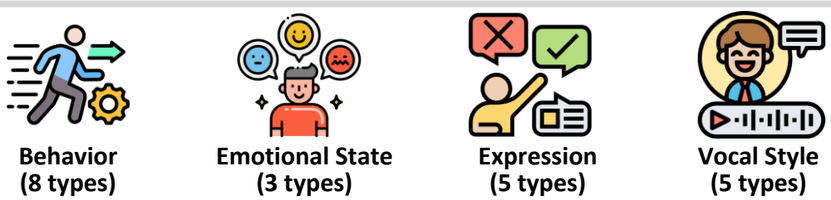
English
512 dialogue turns

Persona and behavior annotation (10 personas and 4 behaviors, 10 times expansion)

Persona Stylization
Big Five Theory

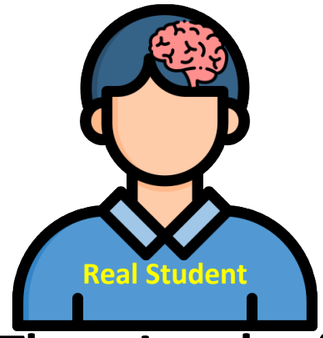


Behavior–Expression Labeling



Step 1. Dataset Construction

Rigorous dataset quality control



Three Levels of Real Students

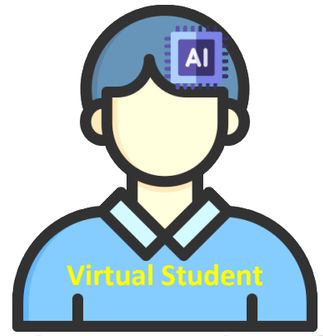
Observable level:
Behaviors, emotions, expressions, and voice align with classroom context.

Inner level:
Responses appear natural and credible, showing real student identity and following classroom norms.

Long-term level:
Maintain target persona maintain a stable persona style throughout extended interaction

Define a hierarchical evaluation framework

Verify whether complex abilities are successfully modeled



Three Tasks of Virutal Students

Task 1 Basic Coherence:
Can virtual students generate multimodal behaviors aligned with context?

Task 2 Student Realism:
Can virtual students be like real students?

Task 3 Persona Consistency:
Can virtual students maintain stable personas during interactions?

Step 2. Evaluation Framework

3 representative open-source LLMs

Using LoRA to fine-tune various virtual student agents

10 different personas

Task 1
Basic Coherence

Task 2
Student Realism

Task 3
Persona Consistency

Step 3. Systematic Experiments and Analysis