VTE VOCATION, TECHNOLOGY & EDUCATION

THEMATIC PAPER: APPRENTICESHIP

Development of vocational education and training chatbot supported by large language model-based multi-agent system

Yuefan Yu, Zhiqun Zhao*

Institute of Vocational and Adult Education, Beijing Normal University, Beijing 100875, China

ABSTRACT

Vocational education and training (VET) chatbots face issues such as difficulties in competence development support, limitations in teaching decision-making, and defects in competence assessment methods. This study constructed an large language model (LLM)-based VET chatbot and designed a human-computer collaborative teaching model. It applied a quasi-experimental design to conduct a teaching experiment in a course titled Electronic Circuits and CAD Plate Making at a VET institution. Results showed that the LLM-based VET chatbot and its application model could significantly improve students' Level 1 and Level 2 professional competence (functional competence and processual competence according to COMET). It also discusses ideas for future optimization of the LLM-based VET chatbot.

Key words: large language model, multi-agent, chatbots, professional competence

INTRODUCTION

Chatbots based on large language models (LLMs), represented by ChatGPT, have a positive impact on education, including vocational education and training (VET). They can serve as intelligent extensions of technology developers' action intentions and autonomous agents of practical activities, translating abstract teaching intentions into concrete, relevant behaviors, demonstrating agency beyond that of traditional chatbots.

Research on using chatbots as teaching agents in VET has emerged in the last decade. Early studies mainly used rule-based approaches to guide students with predefined knowledge-learning paths. Since then, advances in machine learning have driven continued development in the field, reducing the semantic differences between queries and responses, but not changing the nature of the retrieval-based approach. Until the advent of LLMs, the ability for creative and autonomous generation brought generative approaches to the forefront of research. However, the knowledge-based approach remains the main path for VET chatbot design (Makhlouf *et al.*, 2024). Although some studies recognize that traditional teaching usually only transfers knowledge *via* lectures, and lacks interaction, drills, and personalized feedback, the role of LLMs remains passive and responsive (Chang *et al.*, 2024). This may cause students to become overly dependent on LLMs, gradually turning them into "new types of knowledge receivers", and school teaching may shift from traditional teacher-delivered lectures to "intelligent knowledge-feeding" shallow learning.

The research on LLM-based VET chatbots is still in its infancy, and its construction faces multiple challenges: (1) the main function of chatbots is knowledge question-and-answer or guided learning, which makes it difficult to develop students' professional competence; (2)

***Corresponding Author:**
Zhiqun Zhao, Institute of Vocational and Adult Education, Beijing Normal University, Beijing 100875, China. Email: zhiqunzhao@263.net.
Received: 27 March 2025; Revised: 9 May 2025; Accepted: 15 May 2025
https://doi.org/10.54844/vte.2025.0921

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which allows others to copy and redistribute the material in any medium or format non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

chatbots lack scientifically sound professional competence assessment methods applicable to human-computer dialogue teaching and are unable to provide competence-oriented personalized feedback; (3) LLMs' weak multi-step reasoning skills make it difficult to synthesize instructional situations and vocational education laws to make holistic decisions; (4) LLMs lack professional domain knowledge and are prone to the problem of "illusion" in VET.

In response to the above challenges, this study constructs an LLM-based VET chatbot and its application model for class teaching, based on the theoretical foundations of shaping-oriented, action-oriented (Rauner, 2024), "from novice to expert" competence development logic (Dreyfus, 2004) and scaffolding teaching.

CONSTRUCTION OF LLM-BASED VET CHATBOT

The LLM-based VET chatbot constructed consists of a multi-agent system, proxy server, cloud database, and user interface (Figure 1).

The multi-agent system is supported by a multi-agent framework, which was developed with reference to LangGraph and AutoGen (Wu *et al.*, 2023). The framework treats agents and tools as nodes in a cyclic graph, where interactions between nodes are constrained by the edges and edge logic. At the same time, the framework provides a unique Manager agent that can connect with several subgraphs (basic business flows) to form a composite graph. Manager parallelizes the scheduling of subgraphs *via* coroutines, which enable concurrent execution of multiple agent programs within a single thread by suspending and resuming. It can adjust agents' functions by setting prompts, adjust edge logic by calling predefined logic, perceive scene information by accessing cloud databases and utilizing other sensing tools, and inject information to target agents by inserting messages into message lists. The chatbot has three basic business flows: Intelligent Teacher, Task Analyst, and Competence Evaluator.

Based on professional competence development objectives, the Manager determined the scaffolding strategy of the Intelligent Teacher (supported by ERNIE-4.0-Turbo-8K LLM) across the three competence dimensions—functional, processual, and holistic design (Table 1)—and synthesized the work task and work stage to form a prompt, which required the Intelligent Teacher to take the initiative to start scaffolding teaching. Depending on the stage of work, there are two modes: "planning" and "reflecting". In the planning mode, Intelligent Teacher uses procedural scaffolding in the form of questions to dissect the work, guide students to explore in depth, and gradually expand the space for action. In the reflecting mode, the Intelligent Teacher generates hypothetical questions that run counter to the students' action process and uses procedural scaffolding in the form of questions to continuously induce cognitive conflicts, prompting students to reflect on their actions.

Task Analyst multi-agent system (supported by ERNIE-3.5-8K LLM) consists of a master node and three sub-nodes: Action Space Analyst, Decision Analyst, and Knowledge Retriever. Knowledge Retriever explores in depth the key knowledge required to solve complex tasks through multi-step interleaved retrieval. Action Space Analyst outlines key points and solution strategies for solving work tasks under the objective. Decision Analyst uses conceptual knowledge to analyze the deficiencies in student decision-making under the objective. The results of these analyzes will serve as critical cues for Intelligent Teachers and Competence Evaluators.

Competence Evaluator multi-agent system (supported by ERNIE-3.5-8K LLM) consists of a master node, and three sub-nodes for assessing functional, processual, and holistic design competence. The system is designed to determine how a student handles a work task in the current dialog context to meet the objective. Based on the Few-Shot Chain of Thought, the sub-nodes simulate the human evaluation process, gather sufficient evidence-based proof, and then draw conclusions.

METHODOLOGY

The study applied a posttest-only comparison group quasi-experimental design and recruited two parallel groups in the second year of a VET college. To enhance the effectiveness of the LLM-based VET chatbot, a human-computer collaborative teaching model for work-integrated learning (Figure 2) was constructed and applied in the experimental group. The control group used traditional teaching methods. The experiment lasted for four weeks, and a COMET test was conducted for all students.

RESULTS

The LLM-based VET chatbot and its application model significantly improved students' functional competence and processual competence (Table 2).

CONCLUSION

In this case, students' holistic shaping competence was not improved significantly. Possible reasons include the

Table 1: Intelligent teacher's strategy of generating teaching scaffolds

Objective	Work task	Functional competence	Processual competence	Holistic shaping competence
From novice to advanced beginner	Job-oriented work tasks	Analyze visual observations or other sensory perceptions related to operational behaviors, and help students summarize simple operational rules	Analyze context-independent rules of collaboration between operational behaviors to help students optimize their behaviors to reduce wastes	Analyze universal rules related to industry, society, or the environment, and help students optimize their operational behavior accordingly
From advanced beginner to competent	System-atic working tasks	Analyze important conditions in work contexts that affect operational behaviors, and help students adjust their operational behaviors accordingly	Analyze the relationship between user needs and work processes in each stage of the product life cycle, and help students adjust related actions	Analyze important decision-making conditions in work contexts related to industry, society, or the environment, and help students design action
From competent to proficient	Problem-based special working tasks	Help students synthesize theoretical knowledge and practical experience, make holistic considerations of operation-related contexts, and flexibly adjust operational behaviors	Help students to synthesize the impact of the entire product value chain on the work process, and to regulate their actions in the workplace in an integrated manner	Help students to holistically examine work contexts from an integrated perspective of industry, society and the environment, and to make a holistic shaping for action

Note: It usually takes decades to grow from proficient to expert, thus ignoring that stage.

Table 2: Independent-samples *t*-test results of professional competence

Dimensions	Group	*N*	Mean	SD	*t*	*P*
Functional competence	Experimental	52	6.08	2.88	2.23	0.028
	Control	53	4.92	2.43		
Processual competence	Experimental	52	1.51	1.87	2.16	0.034
	Control	53	0.84	1.26		
Holistic shaping competence	Experimental	52	0.72	1.15	1.43	0.157
	Control	53	0.46	0.64		

SD, standard deviation.

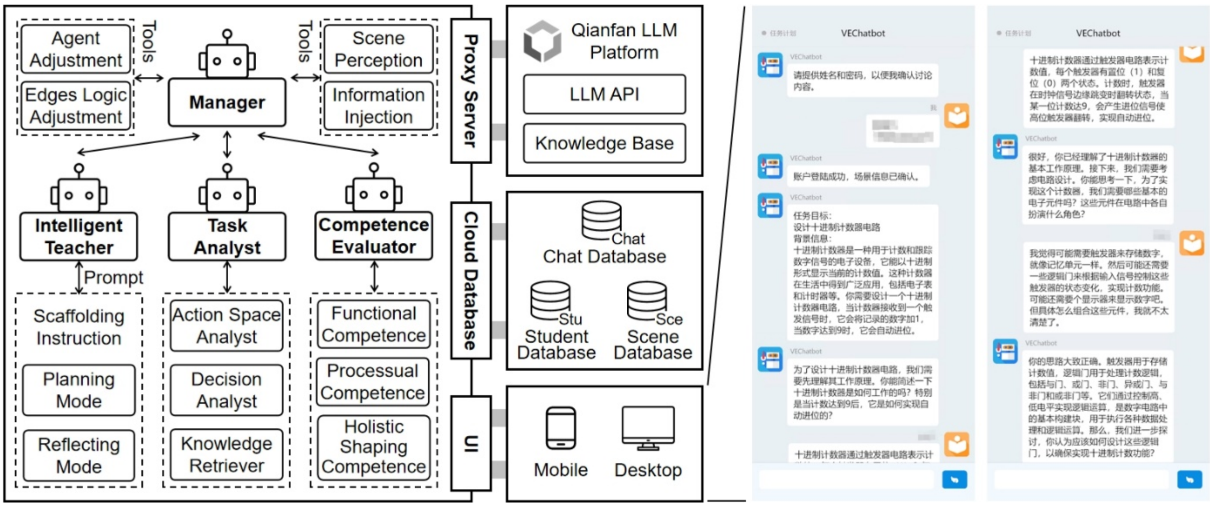

Figure 1. The system architecture of LLM-based VET chatbot. LLM, large language model; VET, vocational education and training; API, application programming interface; UI, user interface.

short duration of the experiment and limited integration of the chatbot with the workplace. The development of professional competence progresses in the order of functional, processual, and holistic shaping. The short experimental period resulted in most students still in the first two stages. Future efforts could extend the experiment duration and embed chatbots into workplace devices. With the help of multi-sensor data, multimodal LLMs can be used to comprehensively perceive both "foreground" and "background" information in real work scenarios. This includes visible state changes of people, machines and objects, as well as implicit information like scene semantics, student emotions, and industry culture. By integrating and reasoning about these temporal features, more contextually appropriate work and teaching strategies can be generated.

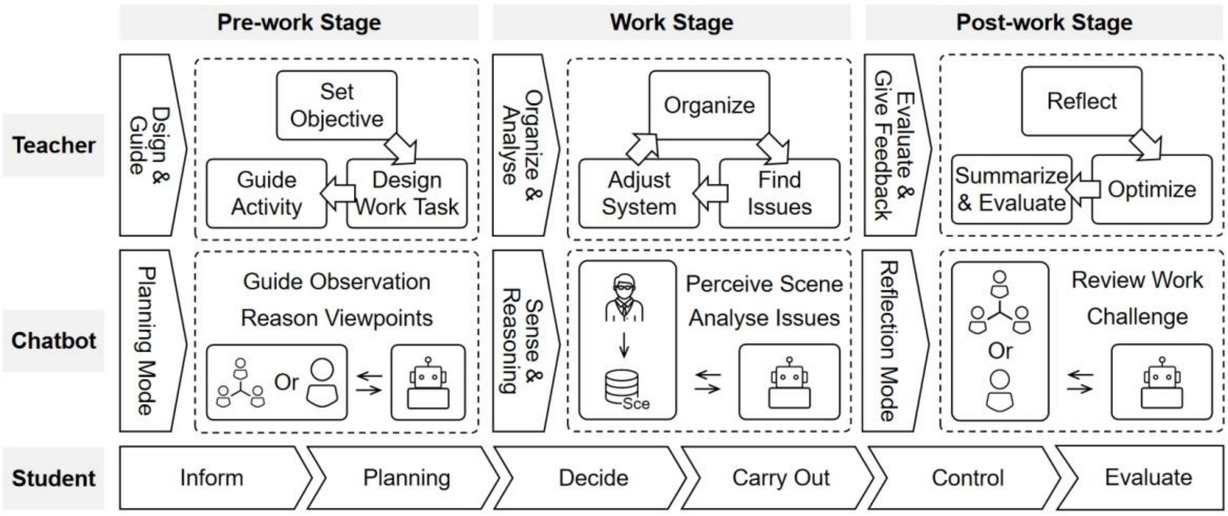

Figure 2. Human-computer collaborative teaching mode for work-integrated learning.

DECLARATIONS

Acknowledgement
None.

Author contributions
Yu YF: Conceptualization, Software, Writing—Original draft. Zhiqun Zhao: Resources, Writing—Review and Editing, Supervision. All authors have read and approved the final version of the manuscript.

Source of funding
This research received no external funding.

Ethical approval
Not applicable.

Informed consent
Not applicable.

Conflict of interest
Zhiqun Zhao is the Associate Editors-in-Chief of the journal. The article was subject to the journal's standard procedures, with peer review handled independently of the editor and the affiliated research groups.

Data availability statement
No additional data.

REFERENCES

Chang, C.Y., Yang, C.L., Jen, H.J., Ogata, H., & Hwang, G. H. (2024). Facilitating nursing and health education by incorporating ChatGPT into learning designs. *Educational Technology & Society*, *27*(1), 215-230.

Dreyfus, H.L. (2004). The five-stage model of adult skill acquisition. *Bulletin of Science Technology & Society*, *24*, 177-181. https://doi.org/10.1177/0270467604264992

Makhlouf, E., Aleneze, A., Shokr, E.A. (2024). Effectiveness of designing a knowledge-based artificial intelligence chatbot system into a nursing training program: a quasi-experimental design. *Nurse Education Today*, *137*, 106159. https://doi.org/10.1016/j.nedt.2024.106159

Rauner, F. (2024). *Handbook of Fundamentals of Modern Vocational Education*. Springer.

Wu, Q. Y., Bansal, G., Zhang, J. Y., Wu, Y. R., Li, B. B., Zhu, E. K., Jiang, L., Zhang, X. Y., Zhang, S. K., Liu, J. L., Awadallah, A. H., White, R. W., Burger, D., & Wang, C. (2023, August 16). AutoGen: Enabling Next-Gen LLM Applications *via* Multi-Agent Conversation. Retrieved March 22, 2025, from https://arxiv.org/abs/2308.08155