Effects of a ChatGPT-based flipped learning guiding approach on learners’ courseware project performances and perceptions

Hai-Feng Li
College of Educational Science, Xinjiang Normal University, China

In recent decades, flipped learning has been adopted by teachers to improve learning achievement. However, it is challenging to provide all students with instant personalised guidance at the same time. To address this gap, based on Chat Generative Pre-trained Transformer (ChatGPT) and the learning scaffolding theory, I developed a ChatGPT-based flipped learning guiding approach (ChatGPT-FLGA) according to the analysis, design, development, implementation and evaluation model. To investigate the effectiveness of ChatGPT-FLGA, a quasi-experiment was conducted in the learning activities of a courseware project. One of two classes was randomly assigned to the experimental group, while the other was assigned to the control group. The students in both classes received flipped classroom instruction and conducted discussions through Tencent QQ applications, but only those in the experimental group learned with ChatGPT-FLGA. The results revealed that the ChatGPT-FLGA significantly improved students’ performance, self-efficacy, learning attitudes, intrinsic motivation and creative thinking. The research findings enrich the literature on ChatGPT in flipped classrooms by addressing the influence of ChatGPT-FLGA on students’ performance and perceptions.

Implications for practice or policy:
• Teachers and universities should utilise ChatGPT as a tool for supporting students’ learning and promoting their problem-solving skills.
• Course designers and academic staff can leverage ChatGPT-FLGA to enact student-centred pedagogical transformation in massive open online courses or flipped learning.
• Course designers should master how to use ChatGPT-FLGA and its learning system, to foster learners’ self-regulated learning, help them promote online self-efficacy and overcome difficulties in learning motivation and creative thinking ability.

Keywords: flipped learning, ChatGPT in education, improving classroom teaching, artificial intelligence learning environments, teaching or learning strategies

Introduction

Multimedia courseware – integrating images, text and media clips in both time and space – has advantages for students’ learning, including diversifying information, improving performances and facilitating learning (Tsai, 2019). To enhance the quality of teaching, an increasing number of teachers design and implement multimedia courseware according to curriculum aims (Kamaruddin, 2010). High-quality multimedia courseware is not only conducive to reflecting the projection relationship, the spatial relationship and the colour of objects but it may also improve the quality of teaching and learning (Yang, 2020).

Therefore, it is crucial for teachers to design high-quality multimedia courseware (Nong, 2016). Pre-service teachers at normal universities, a type of higher education institution in Mainland China specialising in training students to be future educators, need to possess the ability to design multimedia courseware. However, traditional teaching consists mainly of classroom lectures or video-supported teaching (Bartok-Nicolea et al., 2022) and usually involves less time to practise with assistance from the teacher in class (Hsia & Sung, 2020). In view of this, teachers have adopted the flipped learning approach, which reverses the sequence of teachers’ lectures and learners’ assignments (Gutierrez-Gonzalez et al., 2023).
However, in the flipped learning approach, how to provide individual students with instant personalised guidance remains a challenging issue for teachers (Chang & Hwang, 2018). Chat Generative Pre-trained Transformer (ChatGPT), a state-of-the-art artificial intelligence (AI) system developed by OpenAI, can offer a new opportunity to solve this problem. ChatGPT is not only versatile but it can also answer follow-up questions, challenge incorrect premises and admit its mistakes (OpenAI, 2015–2023). ChatGPT has been recognised as an emerging technology assisting students in personalised learning (Z. Zhang et al., 2023). However, few studies have reported how ChatGPT is utilised to facilitate personalised learning in the flipped classroom.

**Literature review**

**Project-based learning for courseware performance**

Teacher lectures and student exercises are the main teaching forms for pre-service teachers in traditional courseware teaching. Traditional teaching may allow learners to be successful with artificial problem sets, but at the cost of decontextualisation and oversimplification (Collins, 1996), which too often fails to prepare students for a great variety of learning and performance required outside of school and denies them the opportunity to take on problem sets in an authentic context. Project-based learning (PBL) has been identified as an effective pedagogy (J. Krajcik et al., 2023) that allows learners to participate in authentic, meaningful work (Collins et al., 1989).

PBL, a form of contextual instruction regarded as an alternative to traditional instruction (C. H. Chen & Yang, 2019), emphasises problem-finding, framing and solving. PBL has six typical characteristics, consisting of a driving question, a focus on learning goals, participation in learning activities, collaboration among learners, the utilisation of scaffolding technologies and the creation of real-world products (J. S. Krajcik & Shin, 2014). The project technique may effectively improve learning performance and thinking (Issa & Khataibeh, 2021) and lead to higher levels of self-reflection and collaboration and higher scores on a standardised test (J. Krajcik et al., 2023).

Providing instant personalised support is crucial for learning activities in PBL. Learners need to work together to solve authentic problems and create tangible artefacts. In this process, teachers and peers, usually as facilitators and collaborators, provide instant feedback and personalised support for students (Guo et al., 2020). Instant feedback and personalised support for learners are conducive to solving authentic problems or making decisions immediately so that skills in creating artefacts and solving problems can continue and develop effectively. Therefore, flipped classrooms are usually integrated with PBL to provide ample opportunities and time for learners.

**Flipped classroom**

The flipped classroom is a popular technology-enhanced teaching approach (Chang & Hwang, 2018) and one of the most promising strategies supporting deeper learning (Cui & Yu, 2019). It focuses more on developing students' higher-order thinking, providing more opportunities and time for learning and permitting self-paced learning (Tavakolizadeh et al., 2012). In the flipped classroom, the roles of teachers and students have undergone significant changes, in which teachers as facilitators assist learners by providing guidance instead of merely delivering learning content, while learners become responsible for their own performance and learning activities.

Flipped classrooms are beneficial in promoting performance. For instance, Gutierrez-Gonzalez et al. (2023) utilised the flipped classroom in a neurosurgery lecture and found that students earned better academic results than in traditional lectures. W. Y. Hwang et al. (2023) implemented a flipped classroom in an engineering education course and found that it contributed to ensuring sustainable learning activities and facilitating cognitive abilities to higher levels. Furthermore, the flipped classroom also shows benefits of improving learners’ perceptions including innovative thinking (Zaina & Martinelli, 2023),
deeper learning (Shen & Chang, 2023); academic self-efficacy and achievement motivation (Bi et al., 2023); critical thinking, problem-solving and creativity (Huang et al., 2022).

As mentioned above, the flipped classroom, also known as the traditional flipped learning guiding approach (T-FLGA), helps promote learning performance and perceptions. However, providing individual students with personalised learning guidance and instant feedback in a project-based activity remains a challenging issue when a teacher needs to interact with students at the same time (Chang & Hwang, 2018). Consequently, it is necessary to integrate technologies into activities in flipped learning to promote instant personalised learning. However, in recent decades, AI chatbots have failed to meet these student needs (H. L. Chen et al., 2020).

**ChatGPT and the flipped classroom**

ChatGPT has four key characteristics: answering follow-up questions, admitting its mistakes, challenging incorrect premises and rejecting inappropriate requests. ChatGPT is not only uncannily adept at mimicking human communication but it is also a versatile wit that can compose essays, translate languages and tell jokes. ChatGPT is already very intelligent. Roivainen (2023) conducted an intelligence quotient (IQ) test for ChatGPT and found that its verbal IQ was superior to that of 99.9% of the test takers. The smartness of ChatGPT has been tested with scholastic assessment test exams, reasoning and legal tests, and has performed better than the human average in many ways (Ray, 2023).

The strengths of ChatGPT may potentially offer opportunities for flipped classrooms based on PBL. ChatGPT may create a ubiquitous learning environment that allows learners to interact with it in out-of-class and in-class activities. ChatGPT may provide appropriate solutions for learners based on individual students' different problems so that personalised learning and guidance can be improved for teaching large classes. The instant and human-mimicking responses have potential strengths, alleviating the problems of students' low level of learning interest and motivation caused by waiting for teacher feedback for a long time. Therefore, this study attempted to integrate ChatGPT into a flipped classroom to improve learners' courseware project performance and perceptions.

**Research questions**

To enhance students’ courseware project performance, I developed a ChatGPT-FLGA by incorporating ChatGPT into flipped learning. I designed and implemented a learning system based on ChatGPT-FLGA, which may provide students with instant personalised guidance, intelligent reminders for learning assignments and guidance collaborative learning. Furthermore, I investigated the effects of ChatGPT-FLGA on students’ courseware project performance and perceptions by exploring the following research questions:

- What is the effect of ChatGPT-FLGA in a flipped learning environment on students’ project performance?
- What is the effect of ChatGPT-FLGA in a flipped learning environment on students’ self-efficacy?
- What is the effect of ChatGPT-FLGA in a flipped learning environment on students’ learning attitudes?
- What is the effect of ChatGPT-FLGA in a flipped learning environment on students’ learning motivation?
- What is the effect of ChatGPT-FLGA in a flipped learning environment on students’ creative thinking?

**Research hypotheses**

Students’ learning experience might be positively affected in a ChatGPT-based learning environment, which enhances their interactions (e.g., active participation, feedback and continuous personalised conversations) (Chan & Hu, 2023). Researchers have pointed out that the features of ChatGPT have many
positive impacts on students' learning performance (Friederichs et al., 2023). Students might resolve study-related issues with the help of ChatGPT, without the need to wait for a long time to discuss with teachers or peers. ChatGPT can lead to increases in learners’ confidence, motivation, engagement and self-efficacy (Wu et al., 2023).

In flipped learning, learners often face difficulties, leading to low engagement, motivation and self-efficacy. ChatGPT may provide more support for students, which is conducive to relieving their learning frustration. Researchers have indicated that the assistance of the chatbot contributes to better self-efficacy, learning attitude and learning motivation (Lee et al., 2022). Chatbot-supported learning has great potential for facilitating learning engagement, learning motivation and learning attitude (H. L. Chen et al., 2020). Students’ higher-order thinking might be positively affected in a ChatGPT-based learning environment (Volante et al., 2023). The key characteristics of ChatGPT are helpful for cultivating students’ divergent thinking (Organisciak et al., 2023).

Based on the literature mentioned above, I believe that ChatGPT-FLGA may provide useful educational value in learning activities in courseware project. Therefore, based on the above research questions, I examined whether ChatGPT-FLGA results in significant differences in learners’ project performances and perceptions compared to T-FLGA. Therefore, from the findings in the literature, I derived five hypotheses that intertwine with one another:

- **H1.** There is significantly greater students’ project performance in the experimental group than in the control group.
- **H2.** There is significantly greater students’ self-efficacy in the experimental group than in the control group.
- **H3.** There are significantly greater students’ learning attitudes in the experimental group than in the control group.
- **H4.** There is significantly greater students’ learning motivation in the experimental group than in the control group.
- **H5.** There is significantly greater students’ creative thinking in the experimental group than in the control group.

**A ChatGPT-FLGA**

**Theoretical background**

In flipped learning, providing adaptive scaffolding of learning with students is a key method for promoting performance. Scaffolding of social constructivism is an influential and related theory (Vygotsky, 1978). Wood et al. (1976, p. 90) defined scaffolding as a process “that enables a child or novice to solve a task or achieve a goal that would be beyond his unassisted efforts”. Vygotsky (1978, p. 86) defined the phenomenon as the zone of proximal development (ZPD), which is “the distance between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem-solving under adult guidance, or in collaboration with more capable peers”.

To assist a student to move through the ZPD, important components are required: a more knowledgeable other, social interactions with a skillful person and scaffolding provided by the more competent person (Karpov, 2005; Vygotsky, 1978). However, human teachers found it difficult to meet these necessary requirements, because they might not provide instant feedback, personalised guidance and appropriate learning scaffold to students (Lajoie, 2005). ChatGPT has the potential to provide learning scaffold and assistance for students. Furthermore, ChatGPT-FLGA needs to proactively provide learning scaffolding based on learners’ abilities and the status of task completion, rather than just being the object of questioning for students.
Instructional design model

The analysis, design, development, implementation and evaluation (ADDIE) model is a generic process that represents a dynamic, flexible guideline for building effective training and performance support tools (Molenda et al., 1996). The ADDIE model has become influential, in producing efficient and effective teaching practices by instructional designers and content developers (Choi et al., 2023). I followed the ADDIE model, described the process of creating ChatGPT-FLGA and the learning system.

In the analysis stage, I aimed to define, identify and determine feasible solutions to problems in T-FLGA. Firstly, I conducted a literature review on the concepts of T-FLGA, its influencing factors and related technologies, and then analysed the key issues faced by teachers and students. Secondly, to investigate the participants’ needs, I interviewed 10 students and three teachers. After completing the data collection, I identified the key issues including scaffolding of learning, instant feedback and conversation mimicking a human. After that, I developed feasible solutions, which consisted of designing and developing ChatGPT-FLGA and the learning system.

In the design and development stage, I implemented the feasible solutions from the analysis stage. Firstly, I designed the learning system based on ChatGPT to meet students’ learning needs. Secondly, to improve the usefulness and ease of use of the learning system, I integrated Tencent QQ and the ChatGPT to form a learning system. Moreover, flipped classrooms need to consider when, where and how to use ChatGPT to provide support for students. During the development phase, I chose technologies for development including the QQ applications, Python programming language and learning analysis.

During the implementation phase, the teacher firstly trained the students in the experimental group to help them master how to use learning system. Then, the students began to learn according to the learning assignment. During the evaluation stage, I examined each step of ChatGPT-FLGA to ensure that they were aligned with the intended goals. The formative evaluation and the summative evaluation were utilised to assess the effects of ChatGPT-FLGA and make timely adjustments and improvements.

ChatGPT-FLGA and the learning system

To promote learners’ courseware project performance, I developed ChatGPT-FLGA and the learning system, as shown in Figure 1. I utilised ChatGPT and QQ to develop the learning system. QQ, an instant messaging software platform in China, is primarily employed for online messaging, voice calls and social networking. In this study, the QQ number based on a QQ proxy server utilised a local server management system to connect to ChatGPT, which achieved instant feedback, personalised communication, continuous conversation and learning intervention.

The system consists of a ChatGPT-based flipped learning system, a QQ learning platform, a database management system and a server management system including ChatGPT, a local server and a QQ server. The local server is responsible for communication with ChatGPT and the QQ servers, processing data and empowering the AI teacher. This system can send prompt requests to the ChatGPT server and then send processed feedback information to the learners. The QQ learning platform is designed and developed based on the QQ function, including an AI teacher, QQ groups, smart reminders, digital materials and learning assignments. The database management system contains conversation databases, assignments databases and learning portfolio databases.
In this study, the learning content is a series of small courseware projects that need to be designed utilising Microsoft PowerPoint, including animation of a single choice and animation of a flying centipede. Students need to log into the QQ groups and view the assignments in the assignment folder, as shown in Figure 2. The assignment content consists of animation videos and learning tasks. The AI teacher will automatically send one of the learning tasks to students in the QQ group every day based on the learning worksheet, to remind students to complete the project on time, as shown in Figure 3. During the pre- and in-class learning process, the learner may continually communicate with the AI teacher by @AI teacher, and the AI teacher can provide the learner with personalised and instant guidance, as shown in Figure 4.
The ChatGPT-based flipped learning system may analyse personal learning participation in the QQ groups with learning analytics and summarize and publish the results. For example, the system can report the status of student posts according to the number of posts within 2 days in the database, as shown in Figure 3.
5. Furthermore, the system may remind learners that the feedback from ChatGPT is not necessarily correct and encourages students to engage in critical conversations with AI teachers through regular information reminders, as shown in Figure 6.

Figure 5. The summary of students’ learning participation provided by the AI teacher

Figure 6. Encouraging students to engage in critical conversations with the AI teacher
Experimental design

Participants

A quasi-experimental design was adopted in this study. A total of 81 participants (aged 20 on average) were from two classes who took a public compulsory course named Modern Educational Technology at a public normal university, representing the average academic level in Mainland China. The participants had no prior knowledge of designing and developing courseware. One intact class of 42 students (18 males) was randomly assigned to be the experimental group, which adopted ChatGPT-FLCA. The other class of 39 students (20 males) was the control group, which adopted T-FLGA. The same teacher instructed both groups. Ethical approval was obtained before the experiment.

Experimental procedure

The experimental procedure is shown in Figure 7. Designing and developing courseware in the Modern Educational Technology course was selected as the learning content for both groups. The experiment began on March 13, 2023. In the first 3 weeks, the teacher introduced knowledge on designing and developing courseware to all students, including theories and operations of Microsoft PowerPoint. In the fourth week, all students completed the pre-test of project performance and the pre-questionnaires. Afterward, the experimental group adopted ChatGPT-FLGA for 3 weeks, while the control group adopted T-FLGA. In the eighth week, all students completed the post-test of project performance and the post-questionnaires. Finally, three students from each group were invited to participate in an interview.

Figure 7. Diagram of the experimental design
Data collection instruments

The learners’ performances of the pre- and post-project were tested with a rubric (J. P. Zhang, 2016, pp. 226–227). The scoring rubric consisted of five dimensions: instruction, science, technology, artistry and operability. Its perfect score was 100, as shown in Table 1. Two experienced instructors who had taught the educational technology course for 4 years scored the two projects. The Kappa’s k of the scoring results of the two projects were 0.64 (p < 0.00) and 0.72 (p < 0.00) respectively, indicating that the scores of the two experts had substantial consistency. The score of the project for each student was calculated by taking the average of the two experts’ scores.

Table 1
Rubric for scoring the project

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5 points</td>
</tr>
<tr>
<td>Instruction</td>
<td>The courseware is not suitable for the needs of teaching objects.</td>
</tr>
<tr>
<td>Science</td>
<td>The courseware has the most contents error.</td>
</tr>
<tr>
<td>Technology</td>
<td>The materials have not been optimised.</td>
</tr>
<tr>
<td>Artistry</td>
<td>The materials’ matching is extremely unreasonable.</td>
</tr>
<tr>
<td>Operability</td>
<td>The courseware does not have a friendly interface.</td>
</tr>
</tbody>
</table>

The Self-efficacy for Learning and Performance Scale developed by Pintrich et al. (1991) was adopted for the study to investigate students’ self-efficacy. This scale consists of eight items, such as “I believe I will receive an excellent grade in this class”. A 5-point Likert scoring scale was adopted, from 5 = strongly agree to 1 = strongly disagree, and its Cronbach’s α value was 0.93, suggesting that the scale is reliable.

The Learning Attitudes Scale developed by G. J. Hwang and Chang (2011) was adopted for the study to investigate students’ learning attitudes. This scale consists of seven items, for example, “I think designing and developing a courseware course is interesting and valuable.” A 5-point Likert scoring scale was adopted, from 5 = strongly agree to 1 = strongly disagree, and its Cronbach’s α value was 0.79, suggesting that the scale is reliable.

The Learning Motivation Scale developed by G. J. Hwang and Chen (2017) was adopted for the study to investigate students’ learning motivation. This scale, consisting of the dimensions of intrinsic and extrinsic motivation, contains six items, such as “If I can, I want to get better grades in this class than most of the other students”. A 5-point Likert scoring scale was adopted and its Cronbach’s α value was 0.79.

The Creative Thinking Tendency Scale developed by Lai and Hwang (2014) was adopted for the study to investigate students’ creative thinking. This scale consists of five items, such as “I like to ask some questions that others do not think of.” A 5-point Likert scoring scale was adopted, from 5 = strongly agree to 1 = strongly disagree, and its Cronbach’s α value was 0.83, implying that the questionnaire is reliable.
The interview questions were modified from the measure developed by G. J. Hwang et al. (2009), which contains seven questions to gain students' perceptions of the learning activity. The interview content was noted down by writing for analysis. Examples of the interview questions include the following: “Overall, what do you think are the advantages of this learning method? Please give an example” and “Would you recommend that teachers use this system or this method for teaching? Why do you think they need this way of teaching? ”

**Experimental results**

One-way covariance (ANCOVA) was utilised to examine students’ performance and perceptions after the experiment by controlling for the effects of the pre-project scores and pre-questionnaire ratings, respectively.

**Project performance**

The homogeneity of the regression coefficient was tested on the two groups’ pre-project performance scores before analysing the students’ project performance by ANCOVA. The assumption of homogeneity of regression was not violated with $F = 2.83$ ($p = 0.10 > 0.05$), indicating that ANCOVA could be employed to analyse the post-project performance of the two groups. As shown in Table 2, the adjusted mean value and standard error were 89.93 and 0.43 for the experimental group and 87.13 and 0.45 for the control group. According to the results ($F = 8.32$, $p = 0.00 < 0.05$), the post-project performance score of the experimental group was significantly higher than that of the control group. The value of partial $\eta^2$ was 0.09, representing a medium effect size (Cohen, 1988).

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>$F$</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>42</td>
<td>88.80</td>
<td>2.73</td>
<td>89.93</td>
<td>0.43</td>
<td>8.32</td>
<td>0.09</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>87.27</td>
<td>3.41</td>
<td>87.13</td>
<td>0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05.

**Analysis of self-efficacy**

The homogeneity of the regression coefficient was tested on the two groups’ self-efficacy ratings. The results showed that the assumption of homogeneity of regression was passed with $F = 1.17$ ($p = 0.28 > 0.05$), revealing that ANCOVA could be used to analyse students’ self-efficacy ratings of the two groups. Table 3 shows the results of the self-efficacy rating for the two groups. The adjusted mean value and standard error were 34.77 and 0.54 for the experimental group and 33.14 and 0.55 for the control group, respectively. According to the results ($F = 4.45$, $p < 0.05$), the self-efficacy rating of the experimental group was significantly higher than that of the control group. The value of partial $\eta^2$ was 0.05, representing a small effect size.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>$F$</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>42</td>
<td>35.05</td>
<td>3.73</td>
<td>34.77</td>
<td>0.54</td>
<td>4.45</td>
<td>0.05</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>32.85</td>
<td>4.54</td>
<td>33.14</td>
<td>0.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05.
Analysis of learning attitudes

The homogeneity of the regression coefficient was tested on the two groups’ learning attitudes ratings. The assumption of homogeneity of regression was not violated with $F = 2.24$ ($p = 0.14 > 0.05$), implying the reasonability of executing ANCOVA to probe the students’ learning attitudes. As shown in Table 4, the adjusted mean value and standard error were 31.59 and 0.52 for the experimental group and 29.91 and 0.53 for the control group, respectively. According to the results ($F = 5.10, p < 0.05$), the self-efficacy rating of the experimental group was significantly higher than that of the control group. The value of partial $\eta^2$ was 0.06, representing a medium effect size.

Table 4
One-way ANCOVA result of students’ learning attitudes

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>F</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>42</td>
<td>31.50</td>
<td>2.76</td>
<td>31.59</td>
<td>0.52</td>
<td>5.10*</td>
<td>0.06</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>30.00</td>
<td>4.21</td>
<td>29.91</td>
<td>0.53</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05.

Analysis of learning motivation

The homogeneity of the regression coefficient was tested on the two groups’ learning motivation ratings. Regression slopes of the two groups’ learning motivation scores are homogenous with $F = 0.09$ ($p = 0.76 > 0.05$), suggesting that ANCOVA could be used to analyse learning motivation ratings of the two groups. As shown in Table 5, the adjusted mean value and standard error were 26.01 and 0.27 for the experimental group and 25.14 and 0.27 for the control group, respectively. The result of ANCOVA showed the learning motivation post-test score of the experimental group was significantly higher than that of the control group ($F = 5.32, p = 0.02 < 0.05$). The value of partial $\eta^2$ was 0.06, representing a medium effect size.

Table 5
One-way ANCOVA result of students’ learning motivation

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>F</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>42</td>
<td>26.17</td>
<td>2.37</td>
<td>26.01</td>
<td>0.27</td>
<td>5.32*</td>
<td>0.06</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>24.98</td>
<td>2.09</td>
<td>25.14</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05.

The homogeneity of the regression coefficient was tested on the two groups’ intrinsic motivation ratings. The assumption of homogeneity of regression slopes was met ($F = 1.07, p = 0.30 > 0.05$), suggesting that ANCOVA could be used to analyse students’ intrinsic motivation ratings of the two groups. The result of the ANCOVA $F$ test (Table 6) indicated the experimental group achieved higher intrinsic motivation ($F = 7.38, p = 0.00 < 0.05$, partial $\eta^2 = 0.09$). The value of partial $\eta^2$ was 0.09, representing a medium effect size.

Table 6
One-way ANCOVA result of students’ intrinsic motivation

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>F</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>42</td>
<td>12.83</td>
<td>1.50</td>
<td>12.74</td>
<td>0.16</td>
<td>7.38*</td>
<td>0.09</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>12.03</td>
<td>1.27</td>
<td>12.12</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.05.
The homogeneity of the regression coefficient was tested on the two groups’ extrinsic motivation ratings. The results showed that the assumptions of homogeneity of regression were not violated with $F = 0.30$ ($p = 0.58 > 0.05$), suggesting that ANCOVA could be used to analyse extrinsic motivation ratings of the two groups. The result of the ANCOVA $F$ test (Table 7) did not indicate a significant difference ($F = 1.01$, $p = 0.32 > 0.05$, partial $\eta^2 = 0.01$).

### Table 7
One-way ANCOVA result of students’ extrinsic motivation

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>F</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>42</td>
<td>13.33</td>
<td>1.48</td>
<td>13.27</td>
<td>0.18</td>
<td>1.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>12.95</td>
<td>1.30</td>
<td>13.02</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Analysis of creative thinking tendency

The homogeneity of the regression coefficient was tested on the two groups’ creative thinking tendency ratings. The test of determining homogeneity of regression was not violated with $F = 3.50$ ($p = 0.07 > 0.05$), suggesting that ANCOVA could be used to analyse creative thinking tendency ratings of the two groups. As shown in Table 8, the adjusted mean value and standard error were 19.07 and 0.44 for the experimental group, and 17.40 and 0.46 for the control group, respectively. According to the results ($F= 6.85$, $p = 0.01 < 0.05$), the creative thinking tendency of the experimental group was significantly higher than that of the control group. The value of partial $\eta^2$ was 0.08, representing a medium effect size.

### Table 8
One-way ANCOVA result of students’ creative thinking tendency

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Adjusted mean</th>
<th>SE</th>
<th>F</th>
<th>Partial $\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental group</td>
<td>42</td>
<td>19.19</td>
<td>3.00</td>
<td>19.07</td>
<td>0.44</td>
<td>6.85*</td>
<td>0.08</td>
</tr>
<tr>
<td>Control group</td>
<td>39</td>
<td>17.28</td>
<td>3.96</td>
<td>17.40</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < 0.05$

### Interview results

The interviews were conducted with six students. The students from the experimental group highly praised the ChatGPT-FLGA. Student 1 mentioned, “The AI teacher was great. I could instantly discuss problems with him”. Student 2 and Student 3 stated, “The AI teacher provided me with good solutions, which expanded my thinking”. However, students from the control group emphasised the advantages of T-FLGA, such as opportunities and time. They also raised many problems. Student 4 and Student 6 expressed, “I had received a response for a long time, so I was very disappointed now”. Furthermore, interviewees expressed a desire to integrate AI teachers into online learning platforms, which would allow them to engage in real-time conversations with the AI teacher while studying online courses.

### Discussion

Students who learned with the ChatGPT-FLGA demonstrated better performance. This finding is in line with that of previous researchers (C. J. Lin & Mubarok, 2021). T-FLGA provided students with videos, courseware and communication to improve their learning (Awidi & Paynter, 2019), but these supports failed to meet the needs of the core practices for PBL, including effective project-based educators, instant feedback and student agency in learning communities (Grossman et al., 2019). In ChatGPT-FLGA, students might use the AI teacher to obtain the relative answers immediately, which exceeded the support of other intelligent chatbots for learning (Hew et al., 2023), elevating students’ performance on the courseware projects (Joo et al., 2000). Furthermore, the conversations in the QQ group also generated new perspectives, which was conducive to further improving the quality and innovation of the courseware.
Students' self-efficacy was significantly improved with the help of the ChatGPT-FLGA. The result is consistent with that of Y. N. Lin et al. (2021) but does not agree with the findings of Fryer et al. (2020), who reported that the AI partner did not predict self-efficacy. The main reason is that the previous AI was "weak AI" that was difficult to meeting personalised learning (Smith & Vickers, 2021); thus, students might lose belief in their own abilities to succeed. ChatGPT might effectively alleviate the weakening of learning interest and frustration caused by waiting for feedback for a long time (Wu et al., 2023). In addition, the human-like robots might make people feel more confident about handling and learning skills (Latikka et al., 2019). ChatGPT was not only uncannily adept at mimicking human communication but also a versatile helper reminding students to think creatively, which also scaffold library, providing learning scaffolding for thinking enlightenment and validation (Dilekçi & Karatay, 2023). The AI teacher based on ChatGPT might be seen as a huge adaptive scaffold library, providing learning scaffolding for students according to their needs. The AI teacher, in assisting students to move through the ZPD, is like a wise person who provided students with learning scaffolding (Vygotsky, 1978; Wood et al., 1976). Furthermore, the AI teacher had the function of reminding students to think creatively, which also contributed to guiding students to think about new ideas.

**Conclusion**

This study provides new valuable insights into how ChatGPT is utilised to support human-computer conversations and how ChatGPT-FLGA promotes students’ performance and perceptions in the flipped classroom. The findings revealed that ChatGPT-FLGA can significantly improve students’ project performance, self-efficacy, learning attitudes, learning motivation and creative thinking tendency, compared with the T-FLGA. The interviewed students’ records provide some evidence that the ChatGPT-based flipped learning system can meet students’ learning needs, such as instant feedback, personal
guidance and intelligent reminders, so that students’ learning interest is stimulated and self-efficacy is promoted, which results in improved learning achievement and learning attitudes.

Limitations and future research

There are some limitations in this study that should be noted. First, ChatGPT-FLGA was developed for the multimedia courseware course. Thus, ChatGPT-FLGA may not have the same effectiveness in other courses. Second, the experimental results may not be generalisable to students from other universities because the experimental university is at the middle level in Mainland China. Furthermore, participants were majoring in chemistry, so it is uncertain whether the experimental results can be generalised to other subjects. In future research, I plan to incorporate strategies for ChatGPT-FLGA to enhance students’ extrinsic motivation and explore students’ performance and perceptions.

Acknowledgements

The author would like to thank the students who participated in this study for their support. This research was funded by the Ministry of Education, China (grant number: DCA180324).

References


**Corresponding author:** Hai-Feng Li, tangshanhf@163.com

**Copyright:** Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence (CC BY-NC-ND 4.0). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.