ISSN: 3008-1629

A Study on the Impact of AI-Assisted Project-Based Learning Design on Innovation Ability

Yingliang Wan[1,*]

[1]School of Data Science, Xianda College of Economics and Humanities, Shanghai International Studies University, Shanghai 200083, China

[*] Corresponding Author

Abstract

This study explores the impact of integrating generative artificial intelligence (AI) into Project-Based Learning (PBL) on the innovation ability of undergraduate students in data science education. Utilizing a quasi-experimental design, we applied AI-empowered PBL to the experimental group and traditional PBL to the control group, comparing the differences between the two groups in computational thinking, data literacy, and creativity. The subjects were undergraduate students majoring in data science at a university in China, and the experiment lasted for one semester. Quantitative data included pre- and post-measurement scales (for computational thinking, data literacy, and innovation self-efficacy) and project outcome scores; qualitative data included interviews, code submissions, and analysis of AI interaction logs. Based on theories such as Vygotsky's Zone of Proximal Development, AI acted as a cognitive scaffold in the experimental group, supporting students in solving tasks beyond their independent capabilities through methods like code auto-completion and dialogue-based Q&A. The research results indicate that PBL integrated with AI can significantly enhance students' higher-order thinking and innovation ability: compared to the control group, the experimental group showed greater increases in problem-solving, critical thinking, and creative thinking abilities, with statistically significant differences; their project works also received higher evaluations in terms of complexity and novelty. Concurrently, qualitative evidence shows that students adeptly used AI tools for flexible exploration and immediate feedback, demonstrating higher engagement and autonomy. However, the study also found that improper use of AI may lead to over-reliance or the spread of bias; therefore, cultivating students' AI literacy is critically important. This research provides an empirical basis and practical framework for effectively integrating AI and PBL in data science education, offering implications for cultivating future-oriented innovative talents.

Keywords

Generative Artificial Intelligence, Project-Based Learning, Innovation Ability, Data Science Education, AI Literacy.

1. Introduction

Cultivating students' innovation ability has become one of the core objectives of higher education in the 21st century. In China, universities are responding to the Ministry of Education's "New Engineering Disciplines" reform initiative, cultivating a new generation of engineering and data science talents through interdisciplinary integration and practical innovation. Project-Based Learning (PBL), as a student-centered, problem-driven teaching strategy, is widely recognized for its contribution to developing students' higher-order thinking skills and innovation literacy. PBL allows students to engage in long-term inquiry and complete

ISSN: 3008-1629

project outcomes in authentic contexts, thus effectively enhancing learning motivation, team collaboration, and creative problem-solving abilities. Meanwhile, the rapid development of artificial intelligence technology is reshaping educational practices. Generative AI (such as GPT-4, GitHub Copilot, etc.) demonstrates powerful capabilities in conversational interaction, content generation, and problem-solving. Introducing AI into teaching is expected to provide students with unprecedented personalized support and cognitive scaffolds: AI systems can answer questions in real-time, diagnose misconceptions, and provide targeted feedback, thereby helping students gain a deeper understanding of complex concepts [1]. Research indicates that the appropriate use of AI tools in the classroom can significantly enhance students' learning performance, for example, by improving conceptual understanding and reducing knowledge misconceptions. At the same time, AI-empowered teaching can also stimulate students' creativity: the integration of AI technology by teachers in instruction can enhance students' creative thinking levels, with student learning engagement playing a mediating role in this process. Although the respective potentials of PBL and AI for cultivating innovation ability have been preliminarily confirmed, there is currently a lack of systematic research on the mechanisms and effects of deeply integrating generative AI into PBL to foster students' innovation ability [2]. This study accordingly raises the question: In undergraduate data science education, can the introduction of AI-supported PBL stimulate students' innovation ability more effectively than traditional PBL? Specifically, how will AI act as a cognitive scaffold during the project-based learning process, influencing the development of students' core competencies such as computational thinking and data-driven problem-solving? This study aims to fill this gap and provide evidence-based guidance for educational practice [3].

2. Literature Review

2.1. Project-Based Learning and Innovation Ability

A large body of research supports the positive effect of PBL on 21st-century skills. PBL originates from constructivist philosophy, emphasizing "learning by doing" in authentic problem contexts, where students construct knowledge through autonomous inquiry and project practice. This pathway is particularly conducive to cultivating creative thinking and innovative problem-solving abilities, as open-ended projects prompt students to propose original insights and reflect on improvements. A meta-analysis integrating experimental studies from the past 20 years indicated that, compared to traditional instruction, PBL significantly enhances learning outcomes such as students' academic achievement, thinking skills, and affective attitudes. Especially in engineering and technology courses, PBL's promoting effect on higher-order thinking skills (such as analysis, synthesis, and innovative thinking) is even more significant. PBL also encourages teamwork and communication, which are equally critical in the innovation process [4]. Therefore, adopting PBL in undergraduate data science education and providing students with interdisciplinary, practical problem tasks is expected to comprehensively exercise their future-oriented innovation literacy [5].

2.2. Artificial Intelligence as a Learning Scaffold

With the maturation of generative AI, the education community has begun to explore its potential for supporting student autonomous learning. Based on Vygotsky's Zone of Proximal Development (ZPD) theory, appropriate guidance and tools can transform students' potential developmental level into their actual developmental level. AI is precisely a new type of intelligent scaffold: for example, Large Language Models (LLMs) can respond instantly to student questions through dialogue, guiding their thought paths, which is similar to a tutor

ISSN: 3008-1629

providing hints and feedback at critical moments [6]. Unlike traditional static resources, generative AI can dynamically adjust explanation methods and difficulty based on different students' responses, providing personalized, real-time instructional support. Researchers are beginning to test the effects of AI scaffolds in practice. For example, an empirical study introduced generative AI support in a programming course, assisting instruction with six stages of scaffolding (from prompt training to AI guidance). The results found that students' scores on assessments of problem-solving, critical thinking, and creativity significantly improved after the course (e.g., effect size for problem-solving ability improvement $d \approx 0.48$, creativity improvement $d \approx 0.46$, both $p < 0.05$). This indicates that the scaffolds provided by AI are beneficial for maintaining and developing students' higher-order thinking skills [7]. Furthermore, analysis of classroom records shows that with AI assistance, students' questions are more targeted, and the feedback they receive is more illuminating, thereby enhancing learning engagement and deep thinking abilities. Of course, the literature also warns of the risks of AI misuse or dependency: if students over-rely on AI for answers, it may weaken their independent thinking abilities [8]. Therefore, some studies emphasize developing students' AI literacy and metacognitive skills to ensure they regard AI as an auxiliary tool rather than a source of answers. In summary, current research provides preliminary evidence for integrating AI into education, pointing out that AI has unique advantages in scaffolding support, personalized feedback, and stimulating creativity. However, in higher education, especially in emerging interdisciplinary fields like data science, how to systematically integrate AI with PBL to maximize the cultivation of students' innovation ability remains to be further explored [9].

2.3. The Specificity of Data Science Education

The data science major requires students to possess multifaceted abilities, including computational thinking, data literacy, and cross-domain problem-solving. These abilities are precisely the concrete embodiment of innovation ability in this field. For example, computational thinking includes problem decomposition, pattern recognition, algorithm design, etc., which is the foundation of data-driven innovation; data literacy involves the ability to acquire, clean, and analyze data and discover insights from it, which is key to proposing innovative solutions. Traditional teaching often imparts skills like programming and statistics in isolation, lacking an environment for integrated application [10]. PBL provides an integration platform, embedding programming, statistics, and domain knowledge into projects that solve real-world data problems. However, beginners often encounter numerous challenges during the project process: complex programming debugging, large-scale data processing, selection of appropriate models, etc [11]. These difficulties, if lacking guidance, may dampen students' creative enthusiasm. Generative AI tools can play a facilitator role here. For example, code generation assistants can accelerate prototype development and reduce tedious coding time, allowing students to focus more energy on creative implementation and problem insight; intelligent conversational assistants can provide hints when students encounter knowledge gaps, such as explaining the principles of an algorithm or comparing the pros and cons of different models; data analysis AI tools can quickly generate visualizations, helping students examine data patterns from new perspectives. These functions are equivalent to equipping students with a "round-the-clock" tutor or teammate, providing support for their exploration of complex topics. Therefore, introducing AI into data science PBL is expected to lower the barrier to innovative practice, enabling students to devote more energy to higher-level thinking and creative attempts. At the same time, students also exercise critical thinking in questioning and evaluating AI outputs during their interactions with AI. This human-AI collaborative learning paradigm fits the practical characteristics of the data science field and is worthy of in-depth research on its pedagogical effects [12].

ISSN: 3008-1629

3. Research Design and Methods

3.1. Research Questions

The core question of this study is: In undergraduate data science education, can AI-supported Project-Based Learning improve students' innovation ability more effectively than traditional Project-Based Learning? To this end, we propose the following specific hypotheses: (1) Compared to PBL without AI, AI-supported PBL can significantly improve students' computational thinking ability, data literacy, and creative problem-solving ability; (2) Students in AI-supported PBL will surpass the control group in the innovation and quality of their project outcomes; (3) The extent of students' effective use of AI tools is positively correlated with the magnitude of improvement in their innovation ability.

3.2. Research Subjects and Course Context

This study selected two parallel classes (approximately 60 students in total) of second-year undergraduate students majoring in data science at a university in Shanghai as subjects. There were no significant differences between the two classes in terms of gender, age, or prerequisite course grades, making them suitable for comparison groups. The study was embedded in the core professional course "Introduction to Data Science," which covers data collection, cleaning, analysis, and machine learning fundamentals, suitable for comprehensive project practice. The course was taught by the same instructor, and to control for instructional consistency, the instructor followed a unified schedule and syllabus [13].

3.3. Experimental Design

A quasi-experimental non-randomized control group pretest-posttest design was adopted. The two classes were randomly assigned as the experimental group (AI-PBL group) and the control group (traditional PBL group), each with about 30 students. The experimental period was 16 weeks, corresponding to one semester. Students in both groups were required to work in groups (4-5 students per group) to complete a comprehensive data science project spanning the entire semester. The project themes centered on real-world problems, for example: "Using public transportation data to predict urban passenger flow peaks and optimize route design." The project implementation followed the typical PBL process: defining a driving question, solution design, in-progress research, outcome production, and in-class presentation and reflection [14].

Intervention Measures for the Experimental Group (AI-PBL Group): The experimental group was required to actively utilize designated AI tools as support during the PBL process. The main tools included: 1) Programming AI assistants (e.g., GitHub Copilot): embedded in the IDE to provide code auto-completion and error debugging suggestions, accelerating programming implementation; 2) Conversational AI tutors (e.g., ChatGPT-4): obtaining theoretical explanations, algorithm hints, data insights, etc., through question-and-answer dialogues; 3) AI data analysis tools (e.g., automated modeling platforms like DataRobot or AI plugins for Python): used for preliminary data exploration and rapid model experimentation. To standardize use, we conducted AI tool training for the experimental group at the beginning of the course, teaching how to write high-quality prompts to obtain useful responses and how to verify the information provided by AI. Throughout the project, the instructor did not provide ready-made solutions but guided students to make full use of AI to find ideas and solve difficulties. For example, when students tried a new algorithm, they could ask ChatGPT for its principles and sample code; when stuck on a code bug, they could use Copilot's suggestions for debugging; when analyzing data features, they could have AI generate various visualizations for inspiration.

ISSN: 3008-1629

The instructor monitored student-AI interaction logs and provided timely guidance to prevent serious misinformation [15].

Measures for the Control Group (Traditional PBL Group): The control group completed the same project according to the conventional PBL model, without using generative AI tools. However, they could consult textbooks, references, or seek help from the instructor through regular means like Q&A sessions. The instructor provided the control group with the same proportion of guidance time (answering common questions) but without involving AI assistance [16].

3.4. Data Collection Methods

This study collected multivariate data to comprehensively evaluate the teaching effectiveness:
Quantitative Data: Pretests and posttests were administered to both groups in the 1st week (start of the project) and the 16th week (end), respectively. Measurement tools included: ① Computational Thinking Test: using an adapted computational thinking ability scale, covering dimensions such as algorithm design, decomposition, and abstraction, Cronbach's $\alpha > 0.8$; ② Data Literacy Questionnaire: assessing students' self-rated abilities in data acquisition, analysis, interpretation, and decision-making; ③ Innovation Self-Efficacy Scale: for students to self-rate changes in their confidence when solving open-ended problems and proposing new ideas. The above scales all used a 5-point Likert scale and their reliability and validity were verified in a pilot study. In addition, each group submitted a project outcome package at the end of the project (including project report, source code, presentation PPT, etc.). We developed evaluation rubrics to score the project outcomes through blind review. Evaluation dimensions included technical completeness (correctness, code quality), innovation (novelty and uniqueness of the solution), complexity (problem difficulty and comprehensiveness of the solution), presentation and expression, etc. Each outcome was independently scored by two unaware reviewers, and the average score was taken [17].

Qualitative Data: To deeply understand the role of AI in the learning process, we collected and analyzed the following qualitative materials: ① Learning Diaries and Interviews: Students in the experimental group were invited to write weekly project journals, recording their AI usage and feelings each week. After the project, semi-structured interviews were conducted with 10 students from the experimental group to discuss how AI tools influenced their thinking processes, teamwork, and innovative attempts. Several students from the control group were also interviewed to understand their project difficulties and coping strategies. ② Process Data: We saved the code version control records (e.g., Git commit logs) and AI interaction logs (e.g., conversation records with ChatGPT) of the experimental group students. Code logs were used to analyze the problem-solving process (e.g., number of debugging attempts, number of feature iterations); AI conversation logs were used to analyze the evolution of question quality (e.g., changes from simple questions to in-depth discussions) and the impact of AI responses on student decisions [18]. ③ Classroom Observation Records: Researchers used non-participant observation to record the performance of both groups in seminar classes, including the liveliness of group discussions, problem difficulty, teacher-student interaction, etc., to corroborate the quantitative results [19].

3.5. Analysis Methods

Quantitative data were statistically analyzed using SPSS. First, inter-group difference tests were conducted on the pretest scores for each ability to confirm initial homogeneity. The main

ISSN: 3008-1629

comparison focused on the posttest scores and incremental differences between the two groups, using independent samples t-tests and repeated-measures ANOVA to test whether the improvements in computational thinking, data literacy, and innovation self-efficacy in the AI-PBL group were significantly higher than those in the control group. Project outcome scores were compared between groups using the Mann-Whitney U test or t-test (depending on the distribution). Qualitative materials were analyzed using the coding methods of grounded theory: interview and log texts underwent open coding and axial coding to extract themes, such as analyzing how students used AI for problem-solving, the difficulties encountered and strategies to overcome them, and the impact of AI on team division of labor and cooperation. Code logs were analyzed through quantitative metrics (e.g., average commit frequency, number of modifications) combined with content analysis to reveal differences in project development trajectories. We will triangulate the quantitative and qualitative results to form a comprehensive explanation of the effects of AI-supported PBL [20].

4. Organization of the Text

4.1. Quantitative Results

Pretest results showed no statistically significant differences ($p > 0.5$) between the two groups on the computational thinking test, data literacy questionnaire, and innovation self-efficacy, indicating comparable initial levels between the groups. After 16 weeks of project-based learning, all posttest indicators improved, but the experimental group's margin of improvement was significantly greater than that of the control group [21].

Computational Thinking Ability: The experimental group's posttest mean score increased from a pretest score of 60.4 to 78.6, while the control group increased from 59.8 to 70.3. Repeated-measures ANOVA showed a significant group × time interaction effect ($F = 8.45$, $p < 0.01$), indicating that the experimental group outperformed the control group in the improvement of computational thinking. In terms of sub-dimensions, the experimental group showed substantial improvement in algorithm design, problem decomposition, and abstract modeling. According to students, the code suggestions and interactive guidance provided by AI helped them master more skills in algorithm implementation and optimization [22]..

Data Literacy: In terms of data processing and analysis capabilities, the experimental group's posttest mean score was approximately 15% higher than the control group's. Particularly in the item "extracting insights from data," the experimental group was significantly better than the control group ($t(58)=2.70$, $p < 0.01$). This may be due to AI tools accelerating the data exploration process, enabling students to try more analysis methods and gain deeper insights. For example, one experimental group student mentioned in an interview: "When discussing data patterns with ChatGPT, I got some new ideas I hadn't thought of, which inspired us to improve our model features." This reflects that AI helped students broaden their thinking in data analysis [23].

Innovation Self-Efficacy and Creativity: The experimental group students' scores on the innovation self-efficacy scale increased from a pretest average of 3.1 to a posttest average of 4.2 (out of 5), a significantly greater increase than the control group ($p < 0.05$). They were more confident in their ability to propose unique solutions and tackle open-ended topics. In creative thinking tests (e.g., open-ended problem-solving scores), the experimental group also outperformed the control group. For example, in the final project reports, multiple groups in the experimental group proposed original improvement plans (such as optimizing predictive

ISSN: 3008-1629

models by integrating external data sources), whereas such extensions were rare in the control group [24]. Overall, students using AI scaffolds showed a higher willingness and ability for innovative attempts, which aligns with the promotion of student creativity by AI found in other studies.

Project Outcome Quality: The project outcome scores, blind-reviewed by experts, showed that the experimental group's average score was 85.3/100, significantly higher than the control group's 78.5/100 (U test, $p < 0.05$). The difference was particularly prominent in the "solution innovation" dimension: 5 groups (out of 10) in the experimental group were rated as having "significant innovation" in their solutions, while only 2 groups in the control group reached this rating level [25]. Furthermore, the experimental group's project reports were generally more outstanding in terms of in-depth data analysis, result visualization, and aesthetics. Reviewing experts provided feedback that the experimental group's work demonstrated a stronger spirit of interdisciplinary integration and independent exploration. This confirms that AI-supported PBL helps students produce higher-quality, more creative outcomes [25].

Table 1: Descriptive Statistics and Inter-group Difference Tests for Experimental (AI-PBL) and Control (Traditional PBL) Groups (Pretest & Posttest)

Measured Variable	*Group*	*N*	*Pretest (M ± SD)*	*Posttest (M ± SD)*	*Statistical Test*	*p-value*
Computational Thinking Ability (0–100 points)	*Experimental*	*30*	*60.4 ± 5.21*	*78.6 ± 6.10*	*Repeated-measures ANOVA (Group ×Time Interaction Effect)*	
	Control	*30*	*59.8 ± 5.43*	*70.3 ± 5.88*	*F(1, 58) = 8.45*	*p < 0.01***
Data Literacy (1–5 points)(Dimension: Extracting Insights from Data)	*Experimental*	*30*	*3.2 ± 0.65*	*4.3 ± 0.55*	*Independent-samples t-test (Posttest Comparison of the "Extracting Insights from Data" Dimension)*	
	Control	*30*	*3.1 ± 0.68*	*3.7 ± 0.60*	*t(58) = 2.70*	*p < 0.01***
Innovation Self-Efficacy (1–5 points)	*Experimental*	*30*	*3.1 ± 0.70*	*4.2 ± 0.58*	*Repeated-measures ANOVA (Group × Time Interaction Effect)*	
	Control	*30*	*3.0 ± 0.72*	*3.4 ± 0.64*	*F(1, 58) = 5.82*	*p < 0.05**
Project Outcome Quality (0–100 points)	*Experimental*	*30*	*—*	*85.3 ± 4.50*	*Mann–Whitney U Test*	

ISSN: 3008-1629

Measured Variable	*Group*	*N*	*Pretest (M ± SD)*	*Posttest (M ± SD)*	*Statistical Test (Posttest Comparison)*	*p-value*
	Control	*30*	*—*	*78.5 ± 6.20*	*U = 268.5*	*$p < 0.05$**

4.2. Qualitative Results

Through the analysis of interviews and logs from experimental group students, we gained an in-depth understanding of the mechanism of AI in the learning process:

AI Usage Patterns and Learning Processes: The most frequently used AI tools by students were the programming assistant (Copilot) and the conversational assistant (ChatGPT). Early logs showed that many students initially used AI only as a shortcut tool, such as directly requesting code snippets or answers. However, under the guidance of the instructor, students gradually learned to use AI in a strategic manner. For example, they began to ask ChatGPT higher-level questions (e.g., "How to improve the performance of model X?") and engaged in critical thinking and verification based on AI's answers. Code commit records indicated that the experimental group students had significantly more development iterations than the control group. They often engaged in rapid trial-and-error and optimization after receiving AI suggestions, achieving more frequent prototype iterations. This "rapid feedback-adjustment" loop helped to spark creativity and refine solutions. One student wrote in a journal: "AI is like our brainstorming partner; many ideas are not necessarily directly feasible, but they stimulate us to think further." This shows that the diverse suggestions provided by AI, even if imperfect, became catalysts for innovative ideas.

Team Collaboration and Division of Labor: Interviews indicated that AI tools influenced the groups' collaboration methods. Some students worried about the unfairness of AI use, and there were debates within groups at the beginning of the project about whether to use AI. However, upon seeing that AI could reduce repetitive labor (such as automatically generating formatted code, writing initial drafts of report abstracts, etc.), most group members accepted using AI to assist with low-level tasks, provided academic integrity was upheld. This allowed group members to devote more time to high-level thinking. Logs showed that some groups even designated one member to be specifically responsible for interacting with AI, who then shared the obtained information with the team. This new division of labor model improved efficiency but also brought challenges: if the member responsible for interacting with AI lacked proficiency, it could lead to the spread of biased information. Consequently, several groups later adopted a "dual-confirmation" mechanism: at least two members independently sought help from AI on the same problem to cross-validate the consistency of the AI's output. This demonstrates that students developed certain critical AI usage strategies in practice, improving their control over information quality.

Impact of AI Scaffolding on Affect and Motivation: Many experimental group students mentioned that AI tools provided "psychological support" when tackling difficult problems in the later stages of the project. When they hit a bottleneck, trying to discuss it with ChatGPT became a way to relieve stress—even if AI did not necessarily give the correct answer, the dialogue process itself alleviated anxiety and promoted a re-examination of the problem. One student joked: "ChatGPT is on call 24 hours; when our group got stuck during a late-night

ISSN: 3008-1629

discussion, its replies made us feel like we weren't fighting alone." This indicates that AI, to some extent, played the role of a learning companion, enhancing students' persistence and engagement. This also corroborates findings from other studies that AI can increase student learning engagement through instant feedback and personalized help.

Table 2: Summary of Qualitative Thematic Analysis of Experimental Group Student Interviews and Logs

Main Theme	*Subtheme*	*Key Findings and Representative Evidence (Quotes or Log Excerpts)*
1. AI Usage Patterns and Learning Processes	*1.1 From Instrumental Use to Strategic Inquiry*	*Finding: Students' use of AI evolved from seeking direct answers (instrumental use) to engaging in critical thinking and advanced problem-solving (strategic use). Evidence:"Early logs show... students directly requested code snippets or answers. However, under teacher guidance... they began asking higher-level questions (e.g., 'How can we improve the performance of model X?') and engaged in critical evaluation and verification."*
	1.2 Accelerated Iteration and Creativity Stimulation	*Finding: Instant feedback from AI (especially coding assistants) supported more frequent "rapid feedback–adjustment" cycles. The diverse suggestions provided by AI acted as catalysts for creativity. Evidence: "Code submission records indicate... the number of development iterations was significantly higher than that of the control group... AI felt like a brainstorming partner—many ideas were not directly feasible but inspired further thinking."*
2. Team Collaboration and Division of Labor	*2.1 Cognitive Task Stratification and Efficiency Improvement*	*Finding: AI took over lower-level repetitive tasks, allowing teams to focus their efforts on higher-order creative thinking. Evidence: "AI helped reduce repetitive workloads (e.g., auto-generating formatted code, drafting report summaries)... enabling team members to devote more time to higher-level thinking."*
	2.2 New Collaboration Challenges and Coping Strategies (AI Literacy)	*Finding: The introduction of AI created new collaboration challenges (e.g., information bias propagation), which led students to develop new quality control strategies such as cross-verification. Evidence: "If the member responsible for interacting with AI lacked sufficient skill, biased information could spread. Therefore, several teams later adopted a 'dual confirmation' mechanism... to cross-verify AI outputs for consistency."*
3. Emotional and Motivational Impacts of AI Scaffolding	*3.1 Emotional Support and Anxiety Reduction*	*Finding: When facing complex PBL (Project-Based Learning) challenges, AI (particularly conversational assistants) provided immediate psychological support, helping relieve anxiety and stress. Evidence: "Many students in the experimental group mentioned... that AI*

ISSN: 3008-1629

Main Theme	*Subtheme*	*Key Findings and Representative Evidence (Quotes or Log Excerpts)*
		tools offered 'psychological support.' When they encountered bottlenecks, conversing with ChatGPT became a way to relieve stress... alleviating anxiety."
	3.2 Enhanced Peer-Like Role and Learning Engagement	*Finding: The round-the-clock availability of AI allowed it to serve as a "learning companion," increasing students' perseverance and engagement in problem-solving. Evidence: "One student joked: 'ChatGPT is on call 24/7—when our group hit a dead end during late-night discussions, its responses made us feel we weren't fighting alone.'... This enhanced students' persistence and engagement."*

4.3. Chapter Summary

Synthesizing the above results, this study finds that AI-supported PBL has a positive impact on cultivating students' innovation ability at both cognitive and non-cognitive levels: it not only improved hard skill indicators (such as computational thinking test scores) but also enhanced soft innovation motivation and confidence by influencing the learning process (such as providing diverse ideas and emotional support).

5. Discussion

The research results support our initial main hypothesis, indicating that integrating generative AI into project-based learning is an effective pedagogical innovation that can promote the development of key innovation abilities in undergraduate students. This finding aligns with the 21st-century educational demand for cultivating higher-order thinking and innovative talents, and it also echoes the expectations for reforming teaching models in emerging fields like data science under China's "New Engineering Disciplines" initiative.

First, in terms of promoting higher-order cognitive abilities, the significant improvement in the experimental group's computational thinking and data literacy demonstrates that AI tools functioned as cognitive scaffolds. This is consistent with Vygotsky's theory: when students are engaged in tasks within their Zone of Proximal Development, AI provides timely support to pull the problem's difficulty back into a solvable range, thus achieving a "just-in-reach" learning experience. Especially in complex tasks like programming and data analysis, AI lowered some technical barriers, allowing students to focus on logical thinking and solution creativity. Moreover, process data shows that AI did not think for the students, but rather inspired more ideas through interaction—this indicates that the correct use of AI helps to expand rather than replace the breadth of students' thinking. It is worth noting that we observed the experimental group's relative progress in areas like code optimization and solution evaluation was less than in other aspects, which may reflect the current limitations of AI in these advanced tasks. This reminds educators to leverage AI's strengths and avoid its weaknesses, using AI more for auxiliary support in preliminary ideation and tedious work, while still having students personally face challenges in parts requiring deep professional judgment to exercise their critical decision-making abilities.

ISSN: 3008-1629

Second, in terms of stimulating creativity and willingness to innovate, AI-supported PBL demonstrated unique advantages. The significant increase in students' innovation self-efficacy implies they are more confident in their own innovation abilities after completing the project. This may stem from two aspects: first, the accumulation of successful experiences. AI's help enabled students to complete tasks they originally found difficult, and this positive feedback enhanced their self-efficacy; second, diversity of thought. AI often provides imaginative ideas or interdisciplinary perspectives that human teachers can hardly cover. Even if some ideas are immature, they broaden the pathways for students to think about problems. This confirms findings from some studies that AI can promote creative thinking through diversified feedback. However, we must also recognize that the cultivation of creativity is not automatically achieved by AI alone. In our interviews, some students mentioned that over-reliance on the initial ideas given by AI had limited their imagination, and they later realized the need to think outside the box of AI prompts. Therefore, when guiding, teachers should emphasize the reference-only nature of AI outputs, encouraging students to critically absorb AI's suggestions rather than accepting them wholesale. This actually cultivates a higher level of metacognition: students learn to reflect on "what needs to be decided by myself, and what can be aided by AI," thus becoming more autonomous innovative learners.

Third, the impact on the affective and motivational levels should not be overlooked. Innovative activities are often accompanied by frustration and uncertainty, and students can easily lose motivation when facing difficulties. This study found that AI tools, to a certain extent, played a psychological support role—providing immediate responses and a sense of companionship, which reduced the common phenomenon of students falling behind in PBL. Especially for introverted students or those with weaker foundations, AI is a "pressure-free" tutor; they can ask it questions repeatedly without fear of ridicule, thus becoming more proactive in overcoming difficulties. This is consistent with previous reports on AI enhancing learning engagement. Of course, we must be vigilant against the misinformation brought by AI hallucinations and biases. Our qualitative analysis showed that a few students were led astray by plausible-sounding but incorrect answers generated by ChatGPT. This highlights the importance of AI literacy education: teachers should explicitly train students to identify the reliability of AI outputs and encourage teamwork to mutually review information provided by AI. Research by Min Zhou et al. also pointed out that students' AI technical literacy significantly moderates the impact of AI on learning outcomes. Therefore, when promoting AI-supported teaching, instruction to improve students' data literacy and AI discernment abilities should be carried out simultaneously, enabling them to command AI tools rather than be led by them.

From a more macro perspective, this study aligns with the current macro-trend of digital transformation in education. The combination of generative AI and PBL demonstrates a new paradigm of technology-empowered pedagogical change: the teacher's role shifts from a knowledge transmitter to a learning designer and facilitator, while AI becomes a controllable auxiliary tool providing differentiated support for each student. This human-computer collaborative model is expected to cultivate talents who better meet the needs of future society—who are not only adept at using advanced tools to improve efficiency but also maintain critical and creative thinking, without getting lost in technology. The data science undergraduate group focused on in this study is precisely the new force for innovation in the future digital economy and AI era, and their development will directly impact the advancement of related fields.

ISSN: 3008-1629

Limitations and Future Research: This study still has some limitations that need to be overcome in future work. First, the sample size is relatively limited, implemented in only a single course at one institution, which affects the broad applicability of the conclusions. Future research could replicate this study in different universities and across different disciplinary backgrounds to validate the universal effectiveness of the AI-integrated PBL model and explore disciplinary differences. Second, the experimental period of this study was one semester, making it impossible to assess longer-term impacts, such as the sustained performance of these students in subsequent courses or actual innovation projects. Future studies could consider longitudinal tracking to observe the shaping of students' long-term innovation literacy by AI-supported learning experiences. Third, we used a quasi-experimental design that was not fully randomized. Although the pretest demonstrated homogeneity between groups, hidden biases might still exist. Subsequent research could attempt randomized controlled trials or crossover experimental designs to enhance the strength of causal inference. Furthermore, the exploration of AI's mechanism of action relied mainly on interview and log analysis, which involves subjective interpretation. Future work could introduce more objective process data metrics, such as using learning analytics techniques to automatically track students' problem-solving paths, AI usage frequency, and patterns, thereby more finely delineating the impact of AI on learning behaviors.

6. Conclusion

This study is a pioneer in empirically validating the value of integrating generative artificial intelligence with project-based learning in the field of data science higher education, demonstrating that this pedagogical innovation can effectively stimulate students' innovation ability. By comparing AI-supported PBL with traditional PBL, we found that AI, as a learning partner and cognitive scaffold, helped students break through several bottlenecks in autonomous inquiry, showing significant advantages in enhancing computational thinking, data literacy, and creativity. At the same time, the integration of AI also changed students' learning methods and mindsets, enhancing their confidence and engagement. However, we also emphasize that technology is not a panacea; its effectiveness depends on reasonable pedagogical design and students' wise use. Educators must continuously balance the relationship between AI assistance and autonomous learning in practice, ensuring that students genuinely improve their innovation ability rather than becoming dependent on technology. Looking ahead, with the further development of AI technology and educational research, we have reason to believe that "AI+PBL" will become one of the important paradigms for cultivating innovative talents. It aligns with the new era's expectations for higher education: using advanced technology to cultivate students' creative spirit and practical abilities in a personalized way, thereby supplying an endless stream of innovative new forces for the country's digital economy and intelligent society. The findings of this study provide strong support for this paradigm and also offer a reference basis for educators and policymakers in advancing teaching reforms. We call for subsequent research to continue deepening exploration in this field, especially focusing on how to better train teachers to master AI-empowered teaching skills, how to establish effective ethical norms to ensure the responsible use of AI in the classroom, and how to promote AI-supported project-based learning models at different educational stages. Through the joint efforts of all parties, artificial intelligence will surely become a powerful engine for educational innovation, helping us cultivate a more creative and adaptive next generation of talents.

Acknowledgements

We thank all the teachers and students who participated in this research, as well as the AI experimental platform that provided technical support. This research was funded by a project from the College of Data Science, Xian Da College of Economics and Humanities, Shanghai International Studies University. During the research process, we strictly adhered to research ethics, and all participants signed informed consent forms.

References

[1] Suwistika, R., Ibrohim, I., & Susanto, H. (2024). Improving critical thinking and creative thinking skills through POPBL learning in high school student. JPBI (Jurnal Pendidikan Biologi Indonesia), 10(1), 115–122. https://doi.org/10.22219/jpbi.v10i1.30172

[2] Davar, N. F., Dewan, M. A. A., & Zhang, X. (2025). AI chatbots in education: Challenges and opportunities. Information, 16(3), 235. https://doi.org/10.3390/info16030235

[3] Chen, Rensi. "The application of data mining in data analysis." International Conference on Mathematics, Modeling, and Computer Science (MMCS2022). Vol. 12625. SPIE, 2023.

[4] Wang Y. Efficient Adverse Event Forecasting in Clinical Trials via Transformer-Augmented Survival Analysis. Preprints, 2025. DOI: 10.20944/preprints202504.2001.v1.

[5] Xu, E., Wang, W., & Wang, Q. (2023). The effectiveness of collaborative problem solving in promoting students' critical thinking: A meta-analysis based on empirical literature. Humanities and Social Sciences Communications, 10(1), 16. https://doi.org/10.1057/s41599-023-01508-1

[6] Wang, F., Zhou, X., Li, K., Cheung, A. C. K., & Tian, M. (2025). The effects of artificial intelligence-based interactive scaffolding on secondary students' speaking performance, goal setting, self-evaluation, and motivation in informal digital learning of English. Interactive Learning Environments, 1, 1–20. https://doi.org/10.1080/10494820.2025.2470319

[7] Sweller, J. (2024). Cognitive load theory and individual differences. Learning and Individual Differences, 110, 102423. https://doi.org/10.1016/j.lindif.2024.102423

[8] Zhang, S. (2025). Investigating Community Interpreting as a Dialogic and Multimodal Process: Developing a Framework for Understanding the Psycholinguistic Process of English-Chinese Sight Translation

[9] Matazu, S. S. (2024). Impact of AI-blended learning and AI-personalized learning on undergraduate biology students' attitude and performance in climate change education. Anchor University Journal of Science and Technology, 5(1), 83–95. https://doi.org/10.4314/aujst.v5i1.8

[10] Lin, T. (2025). Enterprise AI governance frameworks: A product management approach to balancing innovation and risk. International Research Journal of Management, Engineering, Technology, and Science. https://doi.org/10.56726/IRJMETS67008

[11] Tuaputty, H., Alimudi, S., Irene, I., Latuperissa, L. N., & Donkor, A. K. (2023). Project-based learning using a laboratory approach on learning outcomes and critical thinking in marine biology. JPBI (Jurnal Pendidikan Biologi Indonesia), 9(1), 103–114. https://doi.org/10.22219/jpbi.v9i1.24551

[12] Huang, J., Tian, Z., & Qiu, Y. (2025, July). AI-Enhanced Dynamic Power Grid Simulation for Real-Time Decision-Making. In 2025 4th International Conference on Smart Grids and Energy Systems (SGES) (pp. 15-19). IEEE.

[13] Goldie, J. G. S. (2012). Assessing 21st Century Skills. SAGE Publication. https://us.sagepub.com/en-us/nam/assessing-21st-century-skills/book237748

[14] Hutsalo, L., Skliar, I., Abrosimov, A., Kharchenko, N., & Ordanovska, O. (2024). Strategies for developing critical thinking and problem-based learning in the modern educational environment. Multidisciplinary Science Journal, 6(2), 2024ss0209. https://doi.org/10.31893/multiscience.2024ss0209

ISSN: 3008-1629

[15] Yang, C., & Meihami, H. (2024). A study of computer-assisted communicative competence training methods in cross-cultural English teaching. Applied Mathematics and Nonlinear Sciences, 9(1), 45-63. https://doi.org/10.2478/amns-2024-2895

[16] Huang, J., & Qiu, Y. (2025). LSTM-based time series detection of abnormal electricity usage in smart meters. Preprints. https://doi.org/10.20944/preprints202506.1404.v

[17] Alwazzan, M. S. (2024). Investigating the effectiveness of artificial intelligence chatbots in enhancing digital dialogue skills for students. European Journal of Educational Research, 13(2), 573–584. https://doi.org/10.12973/eu-jer.13.2.573

[18] Wang, Y., & Ling, C. (2025). Controlling attributes of .xpt files generated by R. PharmaSUG 2025 conference proceedings. San Diego, CA.

[19] Ellwood, R., & Abrams, E. (2018). Student's social interaction in inquiry-based science education: How experiences of flow can increase motivation and achievement. Cultural Studies of Science Education, 13(2), 395–427. https://doi.org/10.1007/s11422-016-9769-x

[20] Anderson, T., & Soden, R. (2001). Peer interaction and the learning of critical thinking skills. Psychology Learning & Teaching, 1(1), 37–40. https://doi.org/10.2304/plat.2001.1.1.37

[21] Fang, Z. (2025). Microservice-driven modular low-code platform for accelerating SME digital transformation. Preprints. https://doi.org/10.20944/preprints202506.0279.v1

[22] Sani, S. M., Kurniawati, N., & Nurwanti, D. I. (2017). The use of reflective writing to improve students' writing and critical thinking skills. The Tenth Conference on Applied Linguistics and The Second English Language Teaching and Technology Conference in Collaboration with The First International Conference on Language, Literature, Culture, and Education, 105, 331–335. https://doi.org/10.5220/0007166903310335

[23] Ausubel, D. P. (2000). The acquisition and retention of knowledge: A cognitive view. Springer Netherlands. https://doi.org/10.1007/978-94-015-9454-7

[24] Qi, R. (2025). Interpretable slow-moving inventory forecasting: A hybrid neural network approach with interactive visualization. Preprints. https://doi.org/10.20944/preprints202505.1367.v1

[25] Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. Journal of Cleaner Production, 437(September 2023), 140527. https://doi.org/10.1016/j.jclepro.2023.140527

[26] Wang, Y. (2025, July 8). AI-AugETM: An AI-augmented exposure–toxicity joint modeling framework for personalized dose optimization in early-phase clinical trials. Preprints. https://doi.org/10.20944/preprints202507.0637.v1