

AI赋能工科专业学位研究生实践创新能力培养 ——以航天特色应用型高校为例

张爱卿¹, 李金云², 周云英³

¹北华航天工业学院建筑工程学院, 河北 廊坊

²北华航天工业学院马克思主义学院, 河北 廊坊

³北华航天工业学院研究生教学部, 河北 廊坊

收稿日期: 2026年8月2日; 录用日期: 2026年9月16日; 发布日期: 2026年9月29日

摘 要

针对航天特色应用型高校工科专业学位研究生实践培养中实体资源约束、指导覆盖不足、能力评价单一等痛点, 构建“实践场景赋能-指导过程赋能-质量管控赋能-航天特色融合”四位一体的AI赋能提升路径。以某高校土木水利专业学位研究生为对象, 开展为期8周的准实验研究。结果表明, AI赋能可显著提升实践成果质量与创新产出, 优化导师指导时间结构, 同时存在技术依赖、算法偏见等风险。研究提出相应调控策略与推广路径, 可为同类高校实践教学改革提供参考。

关键词

AI赋能, 专业学位研究生, 实践创新能力, 航天应用型高校, 虚拟仿真

Research on AI-Enabled Cultivation of Practical Innovation Ability of Engineering Professional Degree Postgraduates —A Case Study of Aerospace Characteristic Application-Oriented Universities

Aiqing Zhang¹, Jinyun Li², Yunying Zhou³

¹School of Civil Engineering, North China Institute of Aerospace Engineering, Langfang Hebei

²School of Marxism, North China Institute of Aerospace Engineering, Langfang Hebei

³Department of Graduate Teaching, North China Institute of Aerospace Engineering, Langfang Hebei

Received: August 2, 2026; accepted: September 16, 2026; published: September 29, 2026

文章引用: 张爱卿, 李金云, 周云英. AI 赋能工科专业学位研究生实践创新能力培养[J]. 职业教育发展, 2026, 15(10): 9-19. DOI: 10.12677/ve.2026.1510407

Abstract

Aiming at the pain points such as entity resource constraints, insufficient guidance coverage and single ability evaluation in the practical training of engineering professional degree postgraduates in aerospace characteristic application-oriented universities, this paper constructs a four-in-one AI-enabled improvement path of “practice scenario empowerment, guidance process empowerment, quality control empowerment and aerospace characteristic integration”. A quasi-experimental study lasting 8 weeks is carried out with civil and hydraulic engineering professional degree postgraduates from a university as subjects. The results show that AI empowerment can significantly improve the quality of practical results and innovation output, optimize the time structure of tutor guidance, and there are also risks such as technology dependence and algorithm bias. The study puts forward corresponding regulation strategies and promotion paths, which can provide reference for practical teaching reform in similar universities.

Keywords

AI Empowerment, Professional Degree Postgraduates, Practical Innovation Ability, Aerospace Application-Oriented Universities, Virtual Simulation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

(一) 研究背景

2017 年国务院印发《新一代人工智能发展规划》，明确提出“利用智能技术加快推动人才培养模式、教学方法改革，构建包含智能学习、交互式学习的新型教育体系”，将教育智能化提升至国家战略高度 [1]。2018 年教育部《教育信息化 2.0 行动计划》进一步要求“推动人工智能在教学、管理等方面的全流程应用，构建全方位、全过程、全天候的支撑体系” [2]。在政策驱动与技术演进的双重作用下，AI 赋能教育已从概念倡导阶段迈入深度融合阶段，研究生教育作为高层次人才培养的核心环节，正在成为智能化变革的关键场域。

与此同时，我国专业学位研究生招生规模持续扩大，2020 年以来年均增幅约 8%，实践创新能力培养面临日益严峻的资源约束。对于航天特色应用型高校而言，这一矛盾更为突出：航天型号研制对实践环节的保密性、规范性和质量要求极高，而高校实验室条件、企业联合培养基地的承载能力均存在上限。如何在有限实体资源条件下，借助 AI 技术拓展实践教学时空边界、提升指导的精准度与覆盖率、强化过程质量管控，成为亟待破解的现实课题。

(二) 研究意义

当前，多数高校 AI 赋能研究生教育的研究集中于通用工具应用层面，存在“同质化严重、学科适配不足、评价缺乏量化证据”等问题 [3]。本研究以航天特色应用型高校为场域，以土木水利、航天工程等典型工科专业学位研究生为对象，系统构建 AI 赋能实践创新能力培养路径，并通过准实验设计与多元数据采集提供量化证据。研究的核心价值在于：为“AI 辅助 + 导师主导”模式提供可操作的实现方案；为航天特色院校的 AI 教学改革提供可复制的模板；为 AI 赋能效果评价提供融合主观问卷、客观日志与

统计控制的评估范式。

(三) 研究方法

本研究综合采用文献研究法、问卷调查法、准实验研究法和数据挖掘法。首先通过文献梳理界定核心概念、搭建理论框架；其次编制问卷，面向我校 120 名工科专业学位研究生开展实践培养现状调研；在精准诊断痛点的基础上，构建 AI 赋能路径并设计准实验方案；最后，采集问卷调查数据、实验成绩数据与平台日志数据进行多源验证。

2. 核心概念与理论基础

(一) 核心概念界定

AI 赋能教育：指利用机器学习、自然语言处理、知识图谱、虚拟仿真、学习分析等人工智能技术，系统性地重塑教学场景、优化教学过程、提升教育效能的教育信息化高级形态[4]。本研究中的 AI 赋能特指将 AI 工具和智能平台嵌入专业学位研究生实践教学全流程，形成“人机协同”的教学新范式。

专业学位研究生实践创新能力：包括工程实践操作能力、复杂问题分析与解决能力、技术集成与创新应用能力、工程伦理与职业素养四个维度[5]。其核心特征在于“面向真实工程问题、以创新性解决方案为输出”。

航天应用型人才：满足航天企事业单位一线需求，具备扎实工程基础、熟练专业技能和航天质量意识，能够快速适应航天型号研制、生产、测试各环节的高素质技术人才。其核心能力要求包含“严谨细致”的工程素养和对航天质量管理规范的深度理解。

(二) 理论支撑

1、技术增强学习理论(TEL) Kirkwood 和 Price (2014)在系统回顾近十年 TEL 研究后指出，技术对学习的增强作用并非自动发生，只有当技术应用与学习目标、教学情境和评价方式协同设计时，才能真正提升学习成效[6]。这一框架为本研究提供了顶层设计原则：AI 赋能不能停留在工具层面，必须与实践教学的培养目标深度耦合。

2、教育数据挖掘理论(EDM) Romero 和 Ventura (2020)的综述指出，EDM 的核心价值在于从学习者的行为日志中提取可操作的教学洞察，支持个性化指导与预测干预[7]。本研究借鉴 EDM 方法，将虚拟仿真平台的操作日志(迭代次数、耗时、参数修改频次)作为实践过程和能力发展的客观证据，弥补了传统评价仅依赖结果性分数的不足。

3、建构主义学习理论 Piaget 的建构主义主张学习者在真实情境中通过主动探索和协作建构意义[8]。Kolb 的经验学习圈进一步将学习过程分解为“具体经验→反思观察→抽象概括→主动实验”的循环[9]。AI 赋能的虚拟仿真环境恰好提供了安全的“具体经验”和“主动实验”空间，AI 辅助的方案论证则支持“反思观察”与“抽象概括”的认知加工。

(三) 理论与研究设计的对应逻辑

上述理论共同构成本研究的底层逻辑，与后续赋能路径设计形成明确对应：TEL 理论明确 AI 技术必须与培养目标、教学情境协同设计的顶层原则，指导“四位一体”赋能路径的整体架构；EDM 理论提供过程行为数据采集与分析的方法支撑，支撑质量管控赋能模块与效果验证环节；建构主义与经验学习圈理论指导虚拟仿真场景与实践学习环节的设计，实现从具象经验到抽象认知的转化。

3. 工科专业学位研究生实践培养现状与痛点分析

(一) 调研设计与实施

编制《工科专业学位研究生实践创新能力现状调查问卷》，包含“实践资源保障”“导师指导质

量”“个人 AI 素养”“创新能力自评”四个分量表,共 25 题,采用李克特 5 点计分法,整体 Cronbach’s $\alpha=0.87$,信度良好。问卷面向我校航天工程、土木水利、机械等专业学位研究生发放 120 份,回收有效问卷 108 份,有效回收率 90.0%。样本覆盖研一至研三,其中航天方向 43 人(39.8%),土木方向 37 人(34.3%),其他方向 28 人(25.9%)。

(二) 实践创新能力现状

各维度得分如表 1 所示。“技术集成与创新能力”得分最低($M=2.85, SD=0.82$),得分率仅 57.0%,表明创新维度是当前培养的最大短板。这一发现与航天领域对“跨系统集成”和“原创方案设计”的迫切需求形成了尖锐反差。

Table 1. Self-evaluation scores of all dimensions of practical innovation ability (N = 108)

表 1. 实践创新能力各维度自评得分(N = 108)

维度	均值(M)	标准差(SD)	得分率
工程实践操作能力	3.12	0.75	62.4%
问题分析与解决能力	3.24	0.69	64.8%
技术集成与创新能力	2.85	0.82	57.0%
工程伦理与职业素养	3.58	0.64	71.6%

在 AI 工具使用现状方面,仅 12.3%的学生报告“经常使用”AI 辅助实践工具,48.1%“偶尔使用”,39.6%“从未使用”。进一步以年级为分组变量进行单因素方差分析(ANOVA),年级效应显著($F(2, 105)=9.47, p<0.001$),事后检验显示研三学生 AI 使用频率显著高于研一和研二,但这种差异更多来自科研论文写作场景的驱动,而非实践环节的主动应用。

(三) 实践培养痛点分析

通过多选题统计和开放性问题的编码,识别出排名前五的痛点(见图 1 所示):实体实验预约困难(72.2%)、校外导师指导不足(65.7%)、实践过程无人跟踪反馈(58.3%)、实践评价方式单一(51.9%)、个性化指导缺失(47.2%)。这五项痛点的选择率均接近或超过 50%,反映出问题的普遍性。

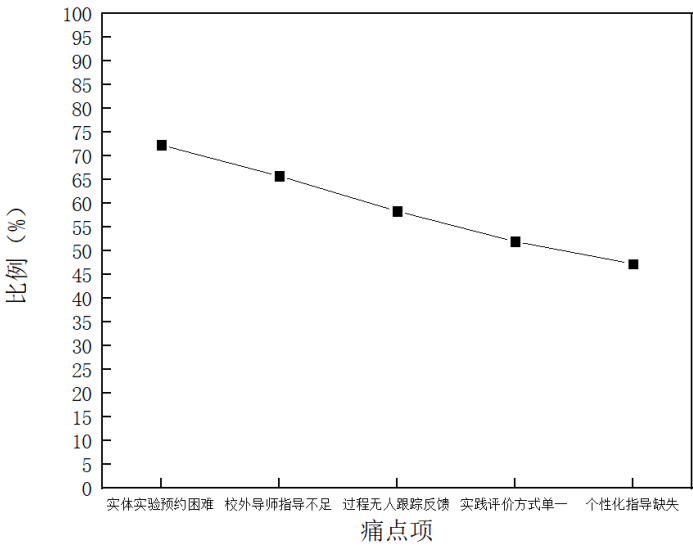


Figure 1. Distribution of main pain points in practical training for engineering professional degree postgraduates (N = 108)

图 1. 工科专业学位研究生实践培养主要痛点分布(N = 108)

从框架分析的角度看,这五类痛点可归入“资源约束-指导短缺-管控缺失-评价单一”四大系统性问题,构成一个相互强化的困局:资源不足限制了实践机会,指导短缺加剧了学生的无效摸索,管控缺失使得问题积累而不被发现,评价单一则无法形成有效的改进反馈。AI技术的介入价值,正在于对这一系统性困局的整体突破。

(四) AI 技术介入的适配性与价值分析

针对上述痛点, AI 技术具有明确的适配路径:

实体资源约束的适配: 虚拟仿真、数字孪生技术可构建与真实工况高度近似的虚拟实践环境,突破场地、设备、安全的物理限制。尤其对于航天领域的高危工况(如火箭发动机点火、月面着陆冲击),虚拟仿真几乎是唯一安全的教学手段。

指导覆盖不足的适配: 基于大语言模型的智能问答系统可提供 7 × 24 小时的常规问题答疑,将导师从重复性事务中解放;学习分析技术可实现多人并发情境下的过程追踪与预警,使“有限导师关注最需要指导的学生”成为可能。

过程管控缺失的适配: EDM 技术可自动采集操作日志、迭代记录、参数修改等过程数据,构建实践行为的数字孪生档案,使“过程可见、问题可知、干预可行”。

评价方式单一的适配: 多模态学习分析可将过程性行为指标(迭代次数、方案多样性、错误修正速度)纳入评价体系,形成“过程 + 结果”的复合评价范式。

4. AI 赋能实践创新能力提升的路径构建

在理论框架的实践转化层面,本节构建的 AI 赋能路径始终回应着第二章的理论基础:虚拟仿真实验的设计遵循 Kolb 经验学习圈,为学生提供沉浸式的“具体经验”; AI 辅助方案论证和智能答疑引导学生进行“反思观察”与“抽象概括”,将感知经验升华为一般性认知结构;创新方案设计环节则对应“主动实验”,推动学生在新的问题情境中检验和应用所学。同时,平台日志数据的采集和分析直接践行了 EDM 的核心原则——从行为数据中提取教学洞察;而“AI 辅助 + 导师主导”的模式设计则体现了 TEL 理论所强调的“技术必须与教学目标和情境协同设计”这一关键主张。以下具体展开四个维度的路径设计。

(一) 实践场景赋能: AI + 虚拟仿真平台

联合航天科技集团共建“航天建造虚拟仿真实验中心”,开发覆盖岩土工程、智能建造、航天器装配等领域的 12 个典型虚拟实验项目。平台具备以下核心功能。

数字孪生模拟: 基于有限元分析和参数化建模技术,构建结构力学响应、基坑变形、航天器装配精度等物理过程的数字孪生模型,学生可通过调整参数实时观察系统响应。例如“月面基地地下结构开挖支护模拟”项目,学生可在 AI 辅助下动态调整支护方案,系统实时呈现应力分布和位移场。

标准化日志采集: 平台自动记录每次实验的操作序列、参数设置、迭代次数、耗时、错误类型等行为数据,为后述的过程评价和 EDM 分析提供数据基础。

个性化难度适配: 基于学生历史操作数据,平台动态调整任务难度。首次完成率达到 85% 以上的学生自动进入“挑战模式”,首次完成率低于 50% 的学生则进入“夯实模式”,实现因材施教。

2024 年,该平台累计完成虚拟实验 1200 余人次,有效缓解了实体实验室的预约压力。

(二) 指导过程赋能: “AI 辅助 + 导师主导”全周期体系

构建覆盖实践全周期的智能指导链,明确 12 个关键 AI 介入节点(参考李擎等提出的 18 环节赋能框架[5]),具体包括以下阶段。

文献梳理阶段: AI 工具自动生成文献综述初稿与创新点线索,缩短从选题到方案设计的周期。试点

中,实验组学生平均文献调研时间由 12.3 小时降至 5.8 小时。

方案论证阶段:利用生成式 AI 快速生成多种技术方案并输出多准则对比分析表。学生须对 AI 生成的方案进行批判性评估和修改,这一过程本身就是创新思维训练。

过程答疑阶段:部署基于课程知识库的智能问答助手,覆盖常见技术问题的即时解答。智能助手的知识库包含课程讲义、行业标准摘要和常见问题解答,经教师审核后上线。

成果打磨阶段:AI 提供论文结构诊断、专利查新、代码规范性检查等服务,辅助学生提升成果的规范性和可提交性。

值得注意的是,针对航天领域的保密要求,本模式中的 AI 工具全部采用本地化部署方案,核心模型运行在校内服务器上,确保学生实践数据不出校园网络。这一设计既是技术适配,也是航天保密文化的具体体现。

在此基础上,进一步开发了“航天故障案例库智能归零演练模块”。该模块收录航天型号研制中的典型质量问题案例(已脱敏处理),学生以角色扮演方式参与“归零”过程:系统模拟故障现象,学生利用 AI 工具进行故障树分析、根因定位和纠正措施设计,AI 系统自动对标航天“双五条”归零标准输出演练评估报告。这一模块使 AI 赋能从通用的“回答问题”深化为航天特有的“质量能力训练”,直接回应了航天单位的核心关切。

(三) 质量管控赋能:动态追踪与航天标准对标

过程数字孪生档案:依托 EDM 技术,建立实践过程数字孪生档案,自动采集虚拟实验操作时间、错误率、方案迭代次数、交互频率等行为数据,实时生成“实践能力显影图”。导师通过管理端可随时查看每位学生的实践画像,精准定位需要干预的环节。

智能量规:设计 AI 驱动的量规(Rubric),从方案创新性(30%)、技术可行性(30%)、实施规范性(20%)、报告质量(20%)四个维度自动评分。试运行期间,AI 评分与专家评分的皮尔逊相关系数达到 0.81,表明 AI 评分具备较高可信度。

航天归零标准自动对标:将航天产品质量问题“归零”五条标准——“定位准确、机理清楚、问题复现、措施有效、举一反三”——嵌入 AI 质量检查模块。学生在提交创新方案或实验报告时,系统自动对标检查:定位准确指问题描述是否清晰可量化;机理清楚指失效机理或设计缺陷的分析是否有理论或数据支撑;问题复现指实验方案是否具备可重复性;措施有效指提出的改进方案是否有验证性实验或模拟数据支持;举一反三指是否考虑了同类问题的系统性排查。系统输出对标评分雷达图和未达标项清单,学生据此返回修改。这一设计使 AI 赋能超越了“打分排名”的浅层应用,深入到航天质量文化的内核培育。

(四) 航天特色融合:精神引领与品质塑造

将航天系统工程思维和质量管理文化融入 AI 教学资源。开发“航天质量文化”系列微课,通过 AI 数字人讲述钱学森、孙家栋等前辈的科研事例和型号研制中的质量把关故事;在虚拟仿真任务中嵌入航天级验收标准,使学生从“做出来”进阶到“做合格”“做可靠”。在故障归零演练模块中,学生不仅学习技术方法,更潜移默化地理解“质量问题归零”背后“一次做对、永远做对”的质量哲学。

5. 试点验证与效果分析

(一) 试点设计

选取我校 2025 级土木水利专业学位研究生“岩土工程数值分析”实践环节进行试点。该课程共计 8 周,含 24 学时实践教学。

分组与同质性保障:两个平行班共 60 名学生,采用整群随机抽样法,以班级为单位通过随机数字表

将两个平行班随机分配为实验组($n=30$)与对照组($n=30$)。整群随机可避免两组学生交叉接触带来的干预扩散,保障实验内部效度。前测数据检验显示,两组在入学成绩(实验组 $M=78.3$, 对照组 $M=77.9$, $t=0.31$, $p=0.76>0.05$)、先修课程绩点(实验组 $M=3.24$, 对照组 $M=3.19$, $t=0.42$, $p=0.68>0.05$)及前测 AI 素养得分(实验组 $M=2.81$, 对照组 $M=2.78$, $t=0.23$, $p=0.82>0.05$)上均无显著差异,保证了分组同质性。

两组均由同一位导师授课和指导,教学内容、实验大纲和考核要求完全一致。唯一区别在于:实验组全程使用 AI 赋能的虚拟仿真平台、智能指导体系和动态质量追踪;对照组在实体实验室完成相同实验内容,按传统方式接受指导。

(二) 实施流程与数据采集

实验组学生首先完成 AI 素养前测和 AI 工具操作考核(前测),随后进入虚拟仿真平台完成 6 项指定实验和 1 项自主创新实验。过程中, AI 助手记录所有操作日志(迭代次数、耗时、参数修改等)。对照组在实体实验室完成相同实验内容。第 8 周,两组统一提交实践报告,并接受实践能力后测、AI 工具操作考核(后测)与满意度问卷调查。

AI 工具操作考核为本研究设计的客观性评价工具:学生在 30 分钟内完成三项规定任务(如文献智能检索与筛选、基于 AI 的方案对比分析、虚拟仿真参数优化),由系统自动评分,60 分以上为通过。该考核旨在测量学生的 AI 工具实际运用能力,弥补自评量表的可能偏差。

(三) 效果对比分析

(1) 实践成果质量

聘请 3 名校外企业专家采用双盲法独立评分,评分者间组内相关系数 $ICC=0.82$,表明评分结果具有良好的信度。创新方案的认定标准为:在导师评审中,被认定为具备至少一项“技术原理新”或“应用场景新”特征的非重复性方案。“技术原理新”指方案采用了不同于既有课程案例的力学模型、算法或结构形式,具有可辨识的原理创新;“应用场景新”指方案将既有技术迁移至尚未在课程中出现过的工程场景,或针对传统场景提出了未被课程案例覆盖的问题定义。多位学生提交相似方案时,仅最早提交的独立方案计入“创新方案”,后续的相似或重复方案不重复计入。描述统计显示,实验组平均分 85.2 ($SD=8.3$),对照组平均分 79.6 ($SD=10.1$)。

为更精确地剥离干预效应,以前测 AI 素养得分为协变量,进行单因素协方差分析(ANCOVA)。首先检验回归斜率同质性假设:组别 $\times$ 前测 AI 素养交互项不显著($F(1, 56)=0.77$, $p=0.384>0.05$),满足斜率同质性条件,可进行标准 ANCOVA,见表 2 所示。

Table 2. ANCOVA results of practical report scores after controlling for pre-test AI literacy

表 2. 控制前测 AI 素养后的实践报告成绩 ANCOVA 结果

变异来源	平方和	自由度	均方	F 值	p 值	偏 η^2
前测 AI 素养(协变量)	382.4	1	382.4	6.83	0.011	0.107
组别(实验/对照)	314.2	1	314.2	5.61	0.021	0.090
误差	3192.1	57	56.0	—	—	—
校正总计	3888.7	59	—	—	—	—

表 2 表明,在控制前测 AI 素养后,组别效应仍然显著($F(1, 57)=5.61$, $p=0.021<0.05$, $n_p^2=0.090$),属于中等偏上的效应量[10]。经协变量调整后的组均值:实验组 85.31 分(标准误 1.37),对照组 79.49 分(标准误 1.37),调整后差值 5.82 分。这意味着排除前测 AI 素养差异的干扰后, AI 赋能的干预效应仍然

成立且稳健。

Table 3. Comparison of innovation scheme output between the two groups
表 3. 两组创新方案产出对比

创新指标	实验组(n = 30)	对照组(n = 30)	组间差异检验
创新方案总数	8	2	—
人均创新方案数	0.27	0.07	$t = 2.17, p = 0.034$
创新方案占比	26.7%	6.7%	$\chi^2 = 4.33, p = 0.037$
“技术原理新”方案数	5	1	—
“应用场景新”方案数	4	1	—

注：一位学生可同时提交同一方案中的多个创新点，因此“技术原理新”和“应用场景新”案例数量之和可能大于创新方案总数。

两组创新方案产出对比见表 3 所示，独立样本 t 检验显示，实验组人均创新方案数(0.27)显著高于对照组(0.07)， $t = 2.17, p = 0.034 < 0.05$ 。卡方检验进一步确认，实验组创新方案占比(26.7%)显著高于对照组(6.7%)， $t = 4.33, p = 0.037 < 0.05$ 。这表明 AI 赋能不仅在整体成绩层面产生了提升效应，在创新性的“上限”表现上同样发挥了促进作用。

(2) 过程数据

过程数据以虚拟仿真平台行为日志为例，具体对比见表 4 所示。

Table 4. Comparison of behavior logs on the virtual simulation platform
表 4. 虚拟仿真平台行为日志对比

指标	实验组(AI 赋能)	对照组(传统模式)	组间差异
平均方案迭代次数(次/任务)	4.2	2.1	+100.0%
方案生成平均耗时(分钟/任务)	18.6	28.8	-35.4%
参数修改平均频次(次/任务)	23.5	11.8	+99.2%
错误类型多样性(种/任务)	5.1	2.4	+112.5%

值得注意的是，实验组虽然迭代次数更多、错误类型更多样，但方案生成耗时反而更短。这一看似矛盾的组合表明：AI 赋能使学生在单位时间内进行了更多轮的“尝试 - 反馈 - 修正”循环——学生更快地生成初稿，发现不足，主动修改，形成了更高效的试错学习节奏。对比之下，对照组学生尽管总耗时更长，但其行为模式以单次“长链路完成”为主，缺少探索性的多方案比较和参数调整，反映出对失败的规避倾向。

(3) 学生满意度

自编《实践教学满意度调查问卷》，包含“学习资源丰富度”“指导及时性”“能力提升感”“总体满意度”四个维度，共 10 题，两组问卷 Cronbach's α 均 > 0.80 。独立样本 t 检验结果如表 5 所示。

Table 5. Comparison of all dimensions of student satisfaction between the two groups
表 5. 两组学生满意度各维度比较

维度	实验组(M $\pm$ SD)	对照组(M $\pm$ SD)	t 值	p 值	Cohen's d
学习资源丰富度	4.35 $\pm$ 0.58	3.12 $\pm$ 0.76	7.03	< 0.001	1.82

续表

指导及时性	4.28 ± 0.62	3.05 ± 0.83	6.49	<0.001	1.68
能力提升感	4.12 ± 0.71	3.48 ± 0.69	3.46	0.001	0.89
总体满意度	4.20 ± 0.65	3.60 ± 0.78	3.18	0.002	0.82

所有维度的效应量(Cohen’s d)均大于 0.8，属于大效应。学习资源丰富度的组间差异最大，印证了虚拟仿真平台在资源供给方面的突出贡献。

(4) 导师指导负担与深度

通过指导日志统计，表 6 呈现了导师指导活动的时间结构变化。

Table 6. Comparison of weekly guidance time allocation of supervisors (hours/week)
表 6. 导师每周指导时间分配对比(小时/周)

指导类型	赋能前基线	试点期间	变化幅度
事务性答疑(重复性问题)	7.8	3.6	-53.8%
常规作业检查与批改	4.7	4.8	+2.1%
深度方案指导(创新论证)	3.5	5.6	+60.0%
合计	16.0	14.0	-12.5%

导师指导总时长减少 12.5%，但结构显著优化：事务性答疑占比从 48.8%下降至 25.7%，深度方案指导占比从 21.9%上升至 40.0%。对焦点访谈的编码分析显示，导师普遍认为“AI 处理了最耗时的重复性答疑，让我有时间和精力与学生进行实质性的方案讨论”。这一结果印证了“AI 辅助 + 导师主导”模式“降负担、提深度”的双重功效。

(5) 客观 AI 素养与能力迁移

AI 工具操作考核：实验组前测通过率 43.3% (13 人)，后测通过率 90.0% (27 人)，提升 46.7 个百分点；对照组前测通过率 40.0% (12 人)，后测通过率 46.7% (14 人)，仅提升 6.7 个百分点。前后测 × 组别的交互效应在广义估计方程(GEE)模型中显著(Wald $\chi^2 = 14.32, p < 0.001$)，表明 AI 工具操作能力的提升归因于干预而非时间效应。

AI 素养自评：实验组前测均值为 2.81 (SD=0.63)，后测为 3.92 (SD=0.51)，配对样本 t 检验($t = 8.24, p < 0.001$)显著；对照组前后测无显著变化($t = 0.65, p > 0.05$)。自评与客观考核结果的方向高度一致，二者交叉验证增强了结论的可信度。

6. 结论、反思与推广建议

(一) 研究核心结论

本研究在 TEL 和 EDM 理论框架下，系统构建并实证检验了航天特色应用型高校 AI 赋能实践创新能力培养模式，得出以下结论。

第一，当前工科专业学位研究生实践培养面临“资源 - 指导 - 管控 - 评价”四重系统短板，AI 技术提供了全链条的突破可能。虚拟仿真平台有效缓解了实体资源约束，智能答疑系统显著降低了导师事务性负担，行为日志采集实现了过程可视化，多元评价体系丰富了能力评估维度。

第二，基于 ANCOVA 的严格检验表明，在控制前测 AI 素养后，AI 赋能对实践成果质量的提升效应仍然显著(调整后差值 5.82 分， $p = 0.021$)。这一结论不依赖于简单的均值比较，排除了“AI 素养高者受益更多”这一备选解释，证明干预效应具有独立的贡献。

第三, AI 赋能的关键机制在于重塑学习过程而非直接提升分数。平台日志分析显示, 实验组的核心优势体现为“多轮快速试错”: 迭代次数翻倍、错误类型多样化、总耗时反降。这表明 AI 真正改变的是学生的探索策略——从“一次做对”的规避型心态转向“快速试错、持续优化”的工程创新思维。

第四, 航天质量标准的深度嵌入使 AI 赋能从“效率工具”升华为“文化载体”。归零五条标准的自动对标系统, 不仅提高了技术规范性的检查效率, 更重要的是将航天质量哲学由外部的文件要求内化为学生的操作习惯和思维框架。

(二) 实践反思与调控策略

本研究在确认 AI 赋能成效的同时, 也识别出若干值得高度重视的新问题, 并针对性提出可落地的调控策略。

技术依赖倾向: 平台日志和课堂观察显示, 实验组中约有 20%(6 名)的学生出现了明显的“AI 先行”行为模式: 面对新任务时, 第一反应是调用 AI 生成参考方案, 而非独立进行问题分析和概念设计。这一发现与崔宇红和韩龙关于研究生生成式 AI 依赖行为的研究相呼应[11]。调控策略: 明确实践任务中“无 AI 独立思考”环节占比不低于 30%, 要求学生在调用 AI 之前先独立形成初步方案框架, 并以“人机协作记录”的形式留存思考痕迹; 将独立性与协作性共同纳入过程性考核, 占总成绩权重的 20%。

AI 评分的算法偏见与规则博弈: 试运行期间, AI 智能评量规与专家评分的相关系数达到 0.81, 但仍存在系统性偏差: AI 倾向于给“结构完整、术语密集、格式规范”的报告打高分, 而对“论证有深度但表达简练、结构不常规但思路有创见”的报告评分偏低。约有 10%(3 名)实验组学生开始无意识地“反向工程”AI 评分规则。调控策略: 明确 AI 智能评量规得分仅作为参考, 占最终成绩权重不超过 30%, 最终成绩以导师与专家评分为准; 定期对 AI 评分模型进行“抗博弈性”审计, 将“模仿 AI 偏好”的文本模式纳入检测范围。

虚拟仿真的“现实折扣”问题: 虽然虚拟仿真项目对标了真实工程工况, 但多位学生在反思日志中指出, 虚拟环境中“失败没有代价”的特性使他们敢于尝试极端参数组合, 这种“冒险”在实体工程中是不被允许的。调控策略: 在虚拟仿真任务中引入“失败代价模拟”机制, 在关键项目中设置虚拟资源约束, 使每一次试错都有模拟的“经济成本”和“安全后果”, 让学生在安全的环境中体验不安全的决定所带来的逻辑后果。

个体差异的放大效应: 分层分析显示, 前测 AI 素养位于最高三分位的学生, 其成绩提升幅度(平均 +8.6 分)显著大于最低三分位的学生(平均 +3.1 分)。AI 赋能虽然在整体上产生了正向效应, 但也可能拉大了不同起点学生之间的差距。调控策略: 为 AI 素养薄弱的学生提供前置性的“AI 工具基础训练营”, 并在实践过程中设置针对性的技术辅导时段, 确保赋能效应的公平性。

以上反思并非否定 AI 赋能的价值, 恰恰相反, 它们揭示了“赋能”与“约束”之间的永恒张力。高质量的教育数字化改革, 不在于回避这些矛盾, 而在于以动态调整的机制持续回应它们的出现。

(三) 推广思路

本模式可向地方同类航天院校及设有航天相关专业的应用型高校推广。推广路径建议分三步:

第一步(0~3 个月): 平台资源与师资协同。通过虚拟教研室机制, 共享虚拟仿真项目库和 AI 智能问答知识库, 降低资源建设门槛; 组织 AI 教学工具专项培训, 提升导师的数字教学能力。

第二步(3~6 个月): 局部试点与标准适配。各校选取 1~2 个工科专业学位研究生培养环节(如课程设计、专业实践)进行试点, 根据本校学科特色调整虚拟仿真项目内容, 复用“AI 辅助 + 导师主导”模式和归零对标检查框架。

第三步(6~12 个月): 跨校协同与数据共享。建立多校联合的 AI 赋能实践教学数据联盟, 在保护隐私前提下共享去标识化的行为日志和评价数据, 为模型的持续优化提供更大的数据底座。通用的“AI 赋能

框架”(场景-指导-管控-文化四位一体)可直接移植至机械、电子、材料等其他工科专业。

需要特别说明的是,对于涉及航天型号背景的案例和数据,推广时须在校际共享协议中明确脱敏标准和保密要求,确保教学资源的合规使用。

(四) 研究不足与未来展望

本研究存在以下局限:首先,试点仅针对土木水利专业、周期为8周,样本量为60人,结论向其他航天工科专业推广的外推性需要更大样本、更长周期的多轮验证。其次,对照组的实体实验室教学本身可能受到“霍桑效应”的影响,即对照组知道自己作为对比班可能产生补偿性努力,导致组间差异被低估,后续研究可设置“等待名单对照组”进一步验证。再次,本研究未对AI工具的长期使用是否导致“技术依赖”和“基础动手能力弱化”进行追踪,也未系统检验AI评价模型的跨学科泛化能力。

未来研究应纳入1学期以上的纵向观察,测量干预撤销后的能力保持效应和技术依赖的消退或固化趋势;同时应探索AI伦理规范的实践落地机制,如建立“AI使用声明”制度,将伦理要求转化为可操作的学术规范。此外,后续将进一步扩大客观平台日志的采集指标范围,以提供更具说服力的循证证据。

基金项目

北华航天工业学院研究生教育教学改革项目(YJY-2026-03)。

参考文献

- [1] 新一代人工智能发展规划[EB/OL].
http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm, 2017-07-20.
- [2] 教育信息化2.0行动计划[EB/OL].
http://www.moe.gov.cn/srcsite/A16/s3342/201804/t20180425_334188.html, 2018-04-18.
- [3] 张煦琳, 龙建纲, 何茂刚. AI赋能的工科研究生跨学科学习的韧性建构——以采集式学习与无界课堂为驱动[J]. 工程研究——跨学科视野中的工程, 2026, 18(4): 533-545.
- [4] 钟云飞, 李学刚. 人工智能赋能传统工科专业教学改革路径研究[J]. 印刷与数字媒体技术研究, 2025(3): 35-41.
- [5] 李擎, 崔家瑞, 杨旭, 等. AI赋能工科专业实践教学管理的探索[J]. 高等工程教育研究, 2025(3): 54-60.
- [6] Kirkwood, A. and Price, L. (2014) Technology-Enhanced Learning and Teaching in Higher Education: What Is ‘Enhanced’ and How Do We Know? A Critical Literature Review. *Learning, Media and Technology*, **39**, 6-36.
<https://doi.org/10.1080/17439884.2013.770404>
- [7] Romero, C. and Ventura, S. (2020) Educational Data Mining and Learning Analytics: An Updated Survey. *WIREs Data Mining and Knowledge Discovery*, **10**, e1355. <https://doi.org/10.1002/widm.1355>
- [8] 皮亚杰. 发生认识论原理[M]. 王宪钊, 译. 北京: 商务印书馆, 1981.
- [9] Kolb, D.A. (1984) *Experiential Learning: Experience as the Source of Learning and Development*. Prentice-Hall.
- [10] Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. 2nd Edition, Erlbaum.
- [11] 崔宇红, 韩龙. 研究生群体的生成式AI依赖: 影响因素、形成机理与行为表现[J]. 情报理论与实践, 2026, 49(5): 60-70.