

• 护理教育 •
• 论 著 •

基于 Kirkpatrick 模型的 AI 赋能护理本科生科研创新能力评价体系的构建与应用

翟淑娜, 王若维, 张佩, 申春红, 于秀娟, 黄金

摘要:目的 构建适用于护理本科生科研创新能力的评价体系,并评价其初步应用效果。方法 通过文献检索、半结构化访谈及 2 轮专家咨询,构建基于 Kirkpatrick 模型的 AI 赋能护理本科生科研创新能力评价体系。采用方便抽样法对 30 名高校护理本科生实施为期 8 周的 AI 辅助科研创新能力培训,并于培训前后应用所构建的评价体系进行效果评价。结果 最终构建的评价框架包含 4 项一级指标、13 项二级指标和 42 项三级指标。培训后护理本科生科研创新能力总分较培训前显著提升($P<0.05$)。结论 构建的 AI 赋能护理本科生科研创新能力评价体系结构合理、内容全面,在短期应用中具有良好的可操作性,可为护理院校开展 AI 辅助科研教学效果评价提供参考。

关键词: 护理本科生; 人工智能; 科研创新能力; 培训效果; Kirkpatrick 模型; 评价体系; 护理教育

中图分类号: R47; G424 **DOI:** 10.3870/j.issn.1001-4152.2026.16.075

Development and application of an AI-enabled research innovation competency evaluation system for undergraduate nursing students based on the Kirkpatrick model

Zhai Shuna, Wang

Ruowei, Zhang Pei, Shen Chunhong, Yu Xiujuan, Huang Jin. School of Nursing, Shandong Xiehe University, Jinan 250109, China

Abstract: **Objective** To develop an evaluation system for assessing research innovation competency among undergraduate nursing students and to evaluate its preliminary application effectiveness. **Methods** An AI-enabled evaluation system for research innovation competency among undergraduate nursing students was developed based on the Kirkpatrick Model through literature review, semi-structured interviews, and two rounds of expert consultation. Using a convenience sampling method, 30 undergraduate nursing students were recruited to participate in an 8-week AI-assisted research innovation training program. The effectiveness of the training was evaluated before and after the intervention using the newly developed evaluation system. **Results** The final evaluation framework consisted of four first-level indicators, 13 second-level indicators, and 42 third-level indicators. The total score for research innovation competency among undergraduate nursing students increased significantly from the pre-training level ($P<0.05$). **Conclusion** The AI-enabled evaluation system for research innovation competency among undergraduate nursing students demonstrates a well-structured framework and comprehensive content. Its preliminary application indicates good feasibility and practicality, providing a useful reference for evaluating the effectiveness of AI-assisted research education in nursing schools.

Keywords: undergraduate nursing students; artificial intelligence; research innovation competency; training effectiveness; Kirkpatrick Model; evaluation system; nursing education

护理科研能力是护理专业核心能力的重要组成部分,而科研创新能力是在科研能力不断积累与深化基础上更深层次的体现,对于推动护理学科发展和提升护理服务质量具有重要意义。然而,当前护理本科生普遍存在科研基础薄弱、科研方法掌握不足、科研实践机会有限等问题,严重制约其科研创新能力的提升^[1]。近年来,人工智能(Artificial Intelligence, AI)技术在教育领域的应用日益广泛,ChatGPT等大语言模型为科研工作提供了新的辅助手段^[2]。《国务院关于深入实

施“人工智能+”行动的意见》^[3]提出,把 AI 融入教育教学全要素、全过程,创新教育教学新模式,推动育人从知识传授为重向能力提升为本转变,为 AI 赋能科研创新能力培养提供了重要政策支撑。在此背景下,AI 工具可在文献检索与分析、研究设计优化、数据处理与可视化、论文写作指导等科研环节提供支持,有望成为提升护理本科生科研创新能力的有效工具^[4]。但开展 AI 辅助护理科研创新能力培训,并对其培训效果进行科学评价是提升护理本科生培养质量的关键环节。Kirkpatrick 四层次评估模型是国际上广泛应用的培训效果评估框架,该模型可从反应层、学习层、行为层和结果层 4 个维度对培训效果进行系统评估^[5]。目前,国内外关于 AI 辅助护理科研创新能力培训效果评价的研究较少,尚缺乏科学、系统的评价指标体系。因此本研究基于 Kirkpatrick 模型,运用德尔菲法和层次分析法,构建 AI 赋能护理本科生科研创新能力评价体系

作者单位:山东协和学院护理学院(山东 济南,250109)

通信作者:黄金,963975446@qq.com

翟淑娜:女,硕士,讲师,314950192@qq.com

科研项目:山东省本科教学改革项目(M2022022);中国民办教育协会 2025 年度规划课题(CANQN250038);山东协和学院校级课程思政教学改革项目(2024SZ05)

收稿:2026-03-10;修回:2026-05-18

并进行应用,旨在为科学评估 AI 辅助护理科研创新能力培训效果提供测量工具。

1 资料与方法

1.1 一般资料

1.1.1 访谈对象 2025 年 1—2 月,采用目的抽样法对来自山东、湖南省 2 所高校的教师开展半结构化访谈。纳入标准:从事护理教育、护理科研或 AI 教学应用研究 ≥ 8 年;自愿参与本研究。排除标准:研究期间不在岗。当访谈资料分析不再出现新的主题,终止访谈。最终共纳入 12 名教师,女 9 名,男 3 名;年龄 38~52(44.62 $\pm$ 5.86)岁;博士 7 名,硕士 5 名;正高级 5 名,副高级 7 名;工作年限 8~25(15.36 $\pm$ 4.13)年。专业领域为:护理教育 6 名,护理科研 4 名,AI 教学应用研究 2 名。

1.1.2 函询专家 2025 年 3—6 月采用目的抽样法遴选专家。入选标准:①护理学、教育学或计算机科学相关专业副高级及以上职称;②从事护理教育、护理科研或 AI 教学应用研究 ≥ 8 年;③近 3 年主持或参与省部级及以上科研课题;④知情同意并承诺全程参与函询。最终遴选来自山东、湖南、江苏、陕西、四川、河南、湖北、广东省 8 个省份 12 所院校的 20 名专家。女 17 名,男 3 名;年龄 36~58(46.78 $\pm$ 6.75)岁;博士 11 名,硕士 9 名;职称为正高级 12 名,副高级 8 名;工作年限 10~25(17.54 $\pm$ 6.02)年;从事护理教育 8 名,护理科研 7 名,AI 教学应用 5 名。

1.1.3 调查对象 2025 年 7—8 月采用方便抽样方式招募山东省某高校 2023 级护理本科生 30 名为研究对象。纳入标准:①护理专业全日制本科生;②年龄 >18 岁,自愿参加本研究并知情同意。排除标准:①正在接受严重的躯体疾病治疗或近期(近 6 个月)因重大疾病住院;②正在参加其他可能干扰结果的研究(例如心理干预试验、行为干预等);③外出交流、请假、休学等研究期间不在学校。共调查 30 人,其中女 25 人,男 5 人;成绩排名前 30%者 11 人,30%~70%者 15 人,后 30%者 4 人;生源地位于城市 9 人,城镇 13 人,农村 8 人。本研究已通过学校伦理审查委员会审批(LLSC-KY01-2025019)。

1.2 AI 赋能护理本科生科研创新能力评价体系的构建

1.2.1 组建研究团队 研究团队包括课题组核心成员 5 名(3 名护理专任教师、1 名临床护理教师、1 名医学统计教师)及外聘专家顾问 2 名(1 名护理科研专家、1 名 AI 教育应用专家),共 7 名成员。5 名课题组核心成员负责文献检索及梳理、访谈提纲制订、指标体系初始框架制订、专家函询问卷编制及实施等;2 名外聘专家顾问负责明确“科研创新能力”的评价边界,确定评价体系框架。

1.2.2 文献梳理 系统检索中国知网、万方数据知

识服务平台、维普网、PubMed 及 Web of Science 数据库,检索时限为建库至 2024 年 12 月。中文检索词包括:科研创新能力,科研创新能力评价,科研能力,科研能力评价;护理本科生;Kirkpatrick 模型;人工智能教育,AI 辅助科研等。英文检索词:nursing students, research innovation ability, scientific research competency; Kirkpatrick model; AI-assisted education, generative artificial intelligence 等,采用主题词与自由词相结合的方式检索。经题目与摘要筛选、全文阅读及文献质量评价,最终纳入中英文文献共 54 篇,重点提炼与 AI 科研创新认知、研究设计等核心能力、AI 工具规范应用及科研创新产出等相关内容,形成条目池。

1.2.3 半结构化访谈 根据研究目的,团队成员制订访谈提纲,主要包括:①AI 赋能背景下护理本科生“科研创新能力”的内涵,其与“科研能力”的差异;②AI 工具在护理科研创新中的应用及相应能力要求;③现行护理科研能力评价工具在创新性评价方面的不足与改进方向。每次访谈时间 45~60 min,访谈内容经录音转录后采用主题分析法,由 3 名团队成员独立编码、逐条归类,经所有成员讨论协商达成共识,提炼出关键维度及指标条目。

1.2.4 初始评价体系构建 研究团队组织线上小组研讨 3 次。在整合文献梳理与访谈结果的基础上,明确“科研创新能力”的评价边界,其中,反应层界定为学生对 AI 科研认知与满意度;学习层聚焦于文献处理、研究设计、数据分析、论文写作等核心能力的习得情况;行为层考察科研实践参与、工具实际运用及创新思维表现等科研行为转化与创新实践;结果层则关注论文发表、项目获批、竞赛获奖等可量化的科研创新产出与发展潜能。在上述分析基础上,研究团队对所有条目进行合并、拆分与语义优化,并依据 Kirkpatrick 模型的四层结构进行归类,经 3 轮研讨与修订,最终形成初始指标条目集。

1.2.5 专家函询 2025 年 3 月,向专家发放包含指标体系、评分说明及修改意见栏的电子问卷进行第 1 轮函询。指标筛选标准设定为重要性均数 ≥ 3.50 、变异系数 $<0.25^{[6]}$ 。2025 年 5 月将第 1 轮统计结果及调整方案反馈给专家,请其重新评判。2 轮函询均采用微信小程序发放问卷,限期 7 d 回收。

1.3 AI 赋能护理本科生科研创新能力评价体系的应用

1.3.1 问卷编制 将构建的科研创新能力评价指标体系进一步形成基于 Kirkpatrick 评估模型的护理本科生科研创新能力评价问卷,将三级指标全部转化为陈述性问卷条目。例如,“利用 AI 高效检索文献”被转化为“我能够使用 AI 进行高效检索文献”。为避免 AI 使用过程中可能出现的学术不端风险,本研究在指标表述中均强调“AI 辅助”与“批判性使用”原则,

明确 AI 仅作为科研辅助工具,而非替代学术思考或原创研究过程。共划分为 4 个维度,13 个能力要素,42 个条目。所有条目采用 Likert 5 级评分法(1=非常不同意,5=非常同意),总分为条目相加所得。在正式测量前,采用便利抽样法选取 20 名护理本科生进行预调查,以检验问卷的内部一致性信度,总量表 Cronbach's α 系数为 0.923,各维度 Cronbach's α 系数为 0.856~0.901。

1.3.2 教学方法 本研究组建由 6 名教师组成的跨学科教学团队:4 名护理专任教师(均具有护理学硕士及以上学历,其中 2 名具有博士学位,均主持或参与过省部级科研项目,具备丰富的护理科研经历及 AI 辅助教学应用实践经验)、1 名临床护理教师(具有 10 年以上临床护理工作经验,熟悉临床科研需求与转化路径)及 1 名医学统计教师(具有医学统计学硕士学历,长期承担本科生科研方法论教学及数据分析

指导工作)。教学团队于正式培训前开展 2 次集体备课,统一教学目标、教学进度及各模块考核标准,确保教学设计的内在一致性。教学团队于 2025 年 7—8 月开展为期 8 周(4 学时/周,共 32 学时)的“AI 赋能护理科研创新能力专题培训”。培训以 Kirkpatrick 四层次模型为设计框架,采用模块化课程设计,各模块明确教学目标、教学方法及阶段性考核方式,确保各层次培训目标可测可评。培训内容共分 5 个模块,见表 1。培训结束后,于 2025 年 9—10 月设置为期 2 个月的“成果与创新转化”跟进阶段,由 2 名护理专任教师以“一对一导师制”方式对学生产出成果(论文初稿、数据分析报告等)进行个性化指导,重点督促学生将培训期间习得的科研创新能力应用于科研实践,推动其向论文投稿、课题申报、竞赛参与等可量化科研创新成果转化。

表 1 教学培训内容

模块名称	周次(学时)	对应层级	教学目标	教学方法与考核	授课教师
AI 工具与科研伦理基础	第 1 周 (4 学时)	反应层	认识 Deepseek、豆包、Kimi、Chat-GPT 等 AI 科研工具的功能与边界;理解 AI 赋能科研创新的价值;掌握 AI 使用学术规范,区分“AI 辅助创新”与“AI 替代思考”	情境导入法+案例分析;课后布置书面反思作业	护理专任教师(主讲)
文献检索与批判性分析	第 2~3 周 (8 学时)	学习层	运用中国知网、PubMed、Web of Science 等开展文献检索;基于 GRADE 标准批判性评价文献质量;利用 EndNote 软件规范管理文献;批判性分析文献局限,提炼创新研究问题	任务驱动+同伴互评;提交检索策略报告及文献综述逻辑框架报告,教师逐一点评	护理专任教师(主讲)
研究设计与数据分析	第 4~5 周 (8 学时)	学习层	借助 AI 批判性确定研究局限,凝练创新性科研问题;AI 辅助调查工具设计;使用 SPSS 统计软件结合 AI 完成数据清洗、统计分析;独立验证 AI 统计建议	问题情境教学+分组实践;提交含 AI 使用记录与批判性注释的分析报告,教师逐份评阅	医学统计教师(主讲)、护理专任教师(协作)
论文写作与原创性规范表达	第 6~7 周 (8 学时)	学习层	在规范框架下借助 AI 梳理论文逻辑,优化学术表达;掌握论文写作规范;识别 AI 写作中的学术不端风险	示范—实践—反思三步法+学术诚信案例研讨;提交含 AI 使用声明的论文初稿,教师逐份批阅	护理专任教师(主讲、协作各 1 名)
科研行为转化与创新实践	第 8 周 (4 学时)	行为层+结果层	规范使用 AI 工具并批判性验证;基于真实需求提出原创性科研问题并设计研究方案;探索 AI 在护理科研中的新用途	完成“科研问题研究方案设计”并口头汇报,教师从创新价值、方案可行性、AI 使用规范性进行反馈	临床护理教师(主讲)、护理专任教师(协作)

1.3.3 评价方法 于培训开始第 1 天及完成全部培训后第 4 个月,采用本研究构建的问卷对 30 名调查对象进行自我报告式评定。

1.4 统计学方法 采用 SPSS26.0 软件进行数据处理与分析。专家积极系数以问卷回收率表示;专家权威程度通过判断依据系数(Ca)与熟悉程度系数(Cs)计算权威系数 $Cr = (Ca + Cs)/2$, Cr 可接受阈值 $> 0.7^{[7]}$;专家意见协调程度以变异系数(CV)和肯德尔和谐系数(W)表示;运用层次分析法确定各级指标权重,并构建判断矩阵,同时采用方根法计算特征向量,

并进行一致性检验,当一致性比率 $CR < 0.10$ 时,判断矩阵通过一致性检验,所得权重有效^[8];服从正态分布的计量资料采用($\bar{x} \pm s$)表示,培训前后比较采用配对样本 t 检验,检验水准 $\alpha = 0.05$ 。

2 结果

2.1 专家函询结果

2.1.1 专家积极性与权威程度 2 轮函询分别发放问卷 20 份,问卷回收有效率为 100%。第 1 轮函询 16 名专家提出文字修改建议共 42 条;第 2 轮函询 8 名专家提出建议 18 条。专家判断依据系数 Ca 分别为

0.92、0.93,熟悉程度系数 C_s 分别为 0.76、0.78,权威系数 C_r 分别为 0.84、0.86。

2.1.2 专家意见协调程度 2 轮函询肯德尔和谐系数分别为 0.185、0.312,可行性评分的和谐系数分别为 0.172 和 0.298(均 $P < 0.001$)。

2.1.3 2 轮函询结果 第 1 轮函询后,删除 1 项因评分未达标准的三级指标;增加 1 项二级指标及相应的 3 项三级指标;使用意愿(愿意使用 AI 工具辅助文献检索、数据分析、论文写作);新增 3 项三级指标,分别涉及文献管理与标注、数据清洗与预处理、遵守 AI 学术规范等内容,合并表述相近的 2 项指标,修改 5 项指标的文字表述使其更加精准。第 2 轮函询后,根据专家反馈微调 3 项指标的措辞。经过 2 轮函询最终形成包含 4 项一级指标、13 项二级指标、42 项三级指标的评估框架。见表 2。对各层级指标矩阵的一致性比率 CR 均 < 0.10 。

2.2 评价体系的应用情况 培训前后护理本科生科研创新能力得分比较,见表 3。

3 讨论

3.1 评价体系具有科学性 专家的选择对最终指标体系的科学性与合理性具有关键性作用。本研究专家遴选覆盖护理教育、护理科研及人工智能教学应用三大核心领域,专业结构合理、代表性好。专家在护理科研能力培养及 AI 技术应用方面具有丰富的实践经验和较高的理论水平,能够从教学实践、科研方法、技术应用及创新能力培养等多个角度对评价指标提出意见和建议,从而保证了评价体系内容的全面性与科学性。2 轮函询问卷回收率均达 100%,表明专家参与积极性高;专家权威系数分别为 0.84 和 0.86,表明专家对研究主题具有充分的判断依据和熟悉程度;2 轮肯德尔和谐系数分别为 0.185、0.312,说明专家意见趋于一致,协调性较好,咨询结果具有较高的可靠性。本研究的评价体系以 Kirkpatrick 四层次评估模型为理论基础,该模型被广泛应用于各类培训项目的效果评价,其系统性、层次性和可操作性得到充分验证^[9-11]。结合 AI 赋能背景,对护理本科生科研创新能力进行层级划分,形成了从主观感受到客观指标、从知识习得到行为转化的完整评价体系。各维度之间既相对独立又相互衔接,符合能力培养与评价要求,从而确保了体系的科学性。

3.2 评价体系具有适用性

3.2.1 反应层 反应层反映护理本科生的 AI 科研认知与满意度,从 AI 工具的认知、使用意愿、满意度进行评价,体现学生的 AI 接受能力、学习动机水平及初步科研认知能力。本研究结果显示,反应层重要性均值为 4.800 分,权重为 0.220,在一级指标中排序相对靠后。说明反应层虽然是培训效果评价的重要组成部分,但单纯提升学生的认知与满意度不足以直接

促进能力提升,更多是为后续能力发展提供基础性支撑^[12]。培训效果显示,反应层在 4 个层级中提升幅度相对最小,可能与培训前学生对 AI 技术已抱有较高期待有关。尽管如此,使用意愿和 AI 工具认知维度的显著改善($P < 0.05$),说明通过系统培训学生不仅加深了对 AI 工具功能边界和学术规范的理解,也进一步强化了规范使用 AI 的意识,为后续科研实践中合理应用 AI 奠定了认知基础。

3.2.2 学习层 学习层反映 AI 辅助科研核心能力习得,从文献检索与管理、研究设计、数据分析、论文写作与规范表达进行评价,能够整体体现学生的科研方法掌握能力、信息整合能力及学术表达能力。学习层重要性均值为 4.850 分,权重为 0.280,在一级指标中权重最高,说明专家对能力习得的高度重视,表明相较于主观认知与满意度,专家更关注学生在知识与技能层面的实际获得情况,强调科研创新能力培养的核心在于方法掌握与能力提升^[13]。从培训效果来看,文献检索与管理提升最为突出,提示 AI 工具在信息获取与整合方面具有显著优势^[14]。相比之下,研究设计和数据分析维度的提升幅度相对较小,可能与这两项能力的习得周期较长、对统计学基础要求较高,需要更长时间的学习和培训方能有更大幅度地提升^[15]。

3.2.3 行为层 行为层反映科研行为转化与创新实践,从科研实践参与、AI 工具实际应用、创新思维进行评估,能够体现学生将已习得的科研知识与技能转化为实际科研行为的能力。本研究结果显示,行为层重要性均值为 4.750,权重为 0.260,仅次于学习层,也说明专家对科研行为转化与创新实践较为重视,提示相较于单纯的知识掌握,科研创新能力培养更强调能力在实践中的应用与转化^[16]。此外,行为层作为连接学习层与结果层的重要桥梁,不仅能够弥补单纯知识评价的不足,还能够反映科研能力在真实科研需求中的应用,是评价体系中不可或缺的核心层级^[12]。本研究结果还显示,行为层提升幅度较为明显,尤其是 AI 工具实际应用和科研实践参与显著提升,提示以工具应用和实践导向为核心的培训模式,有助于推动学生将已有科研认知转化为具体科研行为^[17-18]。

3.2.4 结果层 结果层通过科研创新产出与发展潜能进行解释,主要评价可量化的学术产出、成果质量和自我创新发展,体现护理本科生科研创新能力的最终表现。结果层重要性均值为 4.850,权重为 0.240,相对低于学习层与行为层,提示专家在关注科研最终产出的同时,更加重视能力形成与行为转化过程,说明高质量科研成果的产生通常依赖于长期的知识积累与实践训练,难以在短期培训中显著体现^[19]。结果层能够从科研产出角度对科研创新能力进行评价,是对学习层与行为层的重要补充,有助于构建从能力形成到成果体现的完整评价体系。

表 2 护理本科生科研创新能力评价体系

指标	重要性		可行性		权重	组合权重
	均分($\bar{x}$)	CV	均分($\bar{x}$)	CV		
L1 反应层(AI 科研认知与满意度)	4.800	0.062	4.800	0.052	0.220	
L1-1 AI 工具认知	4.700	0.072	4.800	0.051	0.350	0.077
L1-1-1 了解主流 AI 工具功能	4.750	0.064	4.450	0.081	0.335	0.026
L1-1-2 理解 AI 在科研中的应用价值	4.650	0.069	4.300	0.079	0.327	0.025
L1-1-3 认识 AI 辅助科研的伦理规范	4.800	0.052	4.550	0.067	0.338	0.026
L1-2 使用意愿	4.750	0.064	4.700	0.073	0.320	0.070
L1-2-1 愿意使用 AI 工具辅助文献检索	4.650	0.061	4.400	0.087	0.329	0.023
L1-2-2 愿意使用 AI 工具辅助数据分析	4.800	0.052	4.450	0.096	0.339	0.024
L1-2-3 愿意使用 AI 工具辅助论文写作	4.700	0.053	4.500	0.084	0.332	0.023
L1-3 满意度	4.650	0.071	4.850	0.046	0.330	0.073
L1-3-1 对 AI 工具易用性满意	4.750	0.054	4.300	0.068	0.332	0.024
L1-3-2 对 AI 工具效率提升满意	4.750	0.063	4.600	0.076	0.332	0.024
L1-3-3 对 AI 教学培训满意	4.800	0.052	4.400	0.076	0.336	0.024
L2 学习层(AI 辅助科研核心能力习得)	4.850	0.048	4.700	0.053	0.280	
L2-1 文献检索与管理	4.700	0.072	4.650	0.071	0.280	0.078
L2-1-1 利用 AI 高效检索文献	4.650	0.063	4.350	0.063	0.250	0.020
L2-1-2 利用 AI 分析文献质量	4.600	0.091	4.550	0.070	0.247	0.019
L2-1-3 利用 AI 批判性辅助梳理文献综述框架	4.700	0.064	4.550	0.098	0.253	0.020
L2-1-4 利用 AI 进行文献管理与标注	4.650	0.071	4.480	0.082	0.250	0.020
L2-2 研究设计	4.600	0.077	4.700	0.064	0.230	0.064
L2-2-1 借助 AI 批判性审视现有研究局限	4.550	0.081	4.400	0.076	0.328	0.021
L2-2-2 利用 AI 辅助研究设计	4.650	0.063	4.550	0.070	0.336	0.022
L2-2-3 利用 AI 设计调查工具	4.650	0.079	4.450	0.085	0.336	0.022
L2-3 数据分析	4.600	0.069	4.550	0.087	0.310	0.087
L2-3-1 利用 AI 选择统计方法	4.600	0.057	4.500	0.088	0.244	0.021
L2-3-2 利用 AI 解释统计结果	4.800	0.062	4.500	0.095	0.255	0.022
L2-3-3 利用 AI 生成数据可视化	4.750	0.072	4.550	0.079	0.252	0.022
L2-3-4 利用 AI 进行数据清洗与预处理	4.700	0.068	4.500	0.086	0.249	0.022
L2-4 论文写作与规范表达	4.600	0.076	4.800	0.052	0.180	0.050
L2-4-1 利用 AI 优化论文结构	4.700	0.073	4.450	0.068	0.337	0.017
L2-4-2 在遵循学术规范基础上利用 AI 优化学术表达	4.550	0.073	4.600	0.083	0.326	0.016
L2-4-3 在原创内容基础上利用 AI 进行结构与表达优化	4.700	0.064	4.550	0.073	0.337	0.017
L3 行为层(科研行为转化与创新实践)	4.750	0.064	4.800	0.051	0.260	
L3-1 科研实践参与	4.600	0.067	4.650	0.079	0.300	0.078
L3-1-1 积极参与科研项目	4.650	0.071	4.500	0.081	0.330	0.026
L3-1-2 主动申请科研课题	4.750	0.064	4.600	0.091	0.337	0.026
L3-1-3 参与学术交流活动	4.700	0.053	4.450	0.072	0.333	0.026
L3-2 AI 工具实际应用	4.700	0.064	4.750	0.054	0.360	0.094
L3-2-1 科研全流程中使用 AI	4.800	0.051	4.550	0.084	0.257	0.024
L3-2-2 灵活选择合适的 AI 工具	4.550	0.081	4.400	0.084	0.243	0.023
L3-2-3 批判性评价 AI 输出结果	4.600	0.076	4.400	0.059	0.246	0.023
L3-2-4 遵守 AI 使用的学术规范	4.750	0.058	4.550	0.072	0.254	0.024
L3-3 创新思维	4.500	0.076	4.700	0.064	0.340	0.088
L3-3-1 提出创新性研究问题	4.650	0.052	4.400	0.084	0.330	0.029
L3-3-2 设计新颖的研究方案	4.700	0.080	4.450	0.085	0.333	0.029
L3-3-3 探索 AI 在护理科研的新用途	4.750	0.064	4.500	0.081	0.337	0.030
L4 结果层(科研创新产出与发展潜能)	4.850	0.046	4.750	0.053	0.240	
L4-1 科研产出	4.550	0.070	4.550	0.094	0.340	0.082
L4-1-1 发表学术论文数量	4.800	0.061	4.400	0.066	0.338	0.028
L4-1-2 获批科研项目数量	4.700	0.064	4.500	0.088	0.331	0.027
L4-1-3 参与创新创业竞赛	4.700	0.054	4.450	0.077	0.331	0.027
L4-2 成果质量	4.600	0.067	4.550	0.096	0.350	0.084
L4-2-1 论文发表期刊层次	4.550	0.073	4.500	0.095	0.327	0.027
L4-2-2 科研成果临床转化潜力	4.700	0.072	4.400	0.076	0.338	0.028
L4-2-3 科研成果获奖情况	4.650	0.072	4.400	0.076	0.335	0.028
L4-3 自我创新发展	4.650	0.072	4.650	0.079	0.310	0.074
L4-3-1 自主科研能力显著提升	4.650	0.072	4.500	0.072	0.332	0.025
L4-3-2 科研信心明显增强	4.650	0.079	4.400	0.092	0.332	0.025
L4-3-3 职业发展竞争力提高	4.700	0.072	4.700	0.072	0.336	0.025

表3 培训前后护理本科生科研创新能力得分比较

分, $\bar{x} \pm s$

时间	人数	AI 科研认知与满意度				AI 辅助科研核心能力习得				
		AI 工具 认知	使用意愿	满意度	反应层 均分	文献检索 与管理	研究设计	数据分析	论文写作与 规范表达	学习层 均分
培训前	30	3.63±0.54	3.67±0.48	3.52±0.42	3.61±0.28	2.26±0.29	3.31±0.41	2.59±0.32	3.17±0.48	2.77±0.15
培训后	30	3.96±0.54	4.00±0.42	3.77±0.50	3.90±0.25	3.19±0.74	3.68±0.48	2.88±0.37	3.64±0.48	3.30±0.30
<i>t</i>		2.193	3.218	2.707	4.809	6.119	9.104	3.138	6.421	10.843
<i>P</i>		0.036	0.003	0.011	<0.001	<0.001	<0.001	0.004	<0.001	<0.001

时间	人数	科研行为转化与创新实践				科研创新产出与发展潜能				总体 均分
		科研实践 参与	AI 工具 实际应用	创新思维	行为层 均分	科研 产出	成果 质量	自我创新 发展	结果层 均分	
培训前	30	3.11±0.57	2.34±0.43	3.20±0.58	2.83±0.29	2.50±0.36	2.64±0.21	2.43±0.37	2.62±0.23	2.93±0.10
培训后	30	3.79±0.71	3.20±0.57	3.83±0.55	3.57±0.46	2.86±0.23	2.82±0.36	3.34±0.37	3.24±0.20	3.48±0.18
<i>t</i>		4.189	6.776	4.691	7.556	5.113	2.075	9.021	11.556	13.912
<i>P</i>		<0.001	<0.001	<0.001	<0.001	<0.001	0.047	<0.001	<0.001	<0.001

4 结论

本研究构建的评价体系结构合理,指标内容较为全面,在护理本科生科研培训中具有一定可操作性,可为护理院校开展 AI 辅助科研教学效果评价提供参考工具。本研究存在以下局限:本研究构建的评价体系仅在护理本科生科研培训中进行了短期应用,但尚未开展多中心、大样本的系统测量学检验。后续需基于本研究构建的评价体系开展多中心、大样本研究,对其信度、效度及区分度进行系统检验;同时在不同类型护理院校和不同年级学生中推广应用,以验证评价框架的普适性和推广价值;还需通过纵向追踪研究,系统评估 AI 辅助科研培训对护理本科生科研创新能力提升及学术发展的长期效果;探索该评价框架在临床护理人员继续教育和科研能力培训中的适用性,为护理人才科研能力的全周期培养提供评价依据。

参考文献:

[1] 杨支兰,晋源媛,赵慧敏,等.研一本互助学习坊在护理专业创新型人才培养中的实践[J].护理学杂志,2024,39(17):84-88.

[2] Elpasiony N M A, Sabek E M, Ibrahim S S M. Chat generative pre-trained transformers era: pros and cons between nursing researchers in Egypt[J]. BMC Nurs, 2025, 24(1): 667.

[3] 中华人民共和国国务院. 国务院关于深入实施“人工智能+”行动的意见[EB/OL]. (2025-08-26)[2026-02-24]. https://www.gov.cn/gongbao/2025/issue_12266/202509/content_7039598.html.

[4] Mishra T, Sutanto E, Rossanti R, et al. Use of large language models as artificial intelligence tools in academic research and publishing among global clinical researchers [J]. Sci Rep, 2024, 14: 31672.

[5] Kirkpatrick D L. Evaluating training programs: the four levels[M]. San Francisco: Berrett-Koehler, 1996: 12-26.

[6] Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique[J]. J Adv Nurs, 2000, 32(4): 1008-1015.

[7] 王春枝,斯琴.德尔菲法中的数据统计处理方法及其应用研究[J].内蒙古财经学院学报(综合版),2011,9(4):92-96.

[8] Saaty T L, Vargas L G. Models, methods, concepts & applications of the analytic hierarchy process [M]. 2nd ed. New York: Springer, 2012: 9-10.

[9] Dorri S, Akbari M, Dorri Sedeh M. Kirkpatrick evaluation model for in-service training on cardiopulmonary resuscitation[J]. Iran J Nurs Midwifery Res, 2016, 21(5): 493-497.

[10] Yang M, Zhang X, Han R, et al. Evaluation of the effectiveness of rehabilitation nursing training for clinical nurses based on the Kirkpatrick model[J]. BMC Nurs, 2025, 24(1): 374.

[11] Huang X, Wang R, Chen J, et al. Kirkpatrick's evaluation of the effect of a nursing innovation team training for clinical nurses [J]. J Nurs Manag, 2022, 30(7): 2165-2175.

[12] Alam M F E, Islam A R M. Exploring the hierarchical framework of the Kirkpatrick model in training evaluation from a developing country[J]. Eval Program Plann, 2025, 113: 102696.

[13] 张璐,梁钊,刘旭妍.创新教学方法推动化工类本科生科研创新能力培养[J].日用化学工业(中英文),2024,54(12):1557-1558.

[14] Fenske R F, Otts J A. Incorporating generative AI to promote inquiry-based learning: comparing elicited AI research assistant to PubMed and CINAHL complete[J]. Med Ref Serv Q, 2024, 43(4): 292-305.

[15] Thiem J, Preetz R, Haberstroh S. How research-based learning affects students' self-rated research competences: evidence from a longitudinal study across disciplines[J]. Stud High Educ, 2023, 48(7): 1039-1051.

[16] 吴翟,史志锋,林沛.本科生科研创新能力培养的探索与实践[J].河南教育(高教),2025(4):13-14.

[17] Sobczuk P, Dziedzic J, Bierezowicz N, et al. Are medical students interested in research? -students' attitudes towards research[J]. Ann Med, 2022, 54(1): 1538-1547.

[18] 周海琴,黄珂瑶,王玲玲,等.专科科研小组模式下的导师制带教对护理本科生科研能力的影响[J].护理学杂志,2020,35(21):78-80.

[19] Beasley H K, Vue Z, McReynolds M R, et al. A review of undergraduate research programs aimed at underrepresented students[J]. STAR Protoc, 2024, 5(2): 102997.

(本文编辑 丁迎春)