

# AI 赋能综合实践活动课程教学评一体化的实践探索

——以高中机器狗校园服务项目为例

刘华伟

(上海市徐汇中学)

**摘要:** 综合实践活动课程因其高度的开放性和过程性,普遍面临“教、学、评”割裂的困境——教师凭经验施教、学生缺支架自学、评价靠印象判断,三者各自为政。本文以高一年级“AI 赋能·智慧校园”机器狗校园服务项目为案例,探索 AI 技术支撑下综合实践活动课程教学评一体化的实施策略。课程围绕“校园顺风狗、赛博舞伴、暗夜巡更犬”三个真实校园服务场景,采用项目式学习方式,通过四条核心策略——真实情境锚定教学评共同起点、结构化工具实现“学即评”的嵌入机制、AI 技术降低门槛释放思维空间、迭代循环驱动深度学习与精准评价——构建了教学评同向运行的实践路径。研究表明,上述策略能够有效弥合综合实践活动课程中教学与评价的断裂,促进学生在问题解决、创意物化、责任担当等维度的发展,为同类课程的教学评一体化实施提供可迁移的经验。

**关键词:** AI 赋能; 教学评一体化; 综合实践活动课程

## 一、问题提出

综合实践活动课程是国家课程方案中的重要组成部分。<sup>①</sup>与学科课程相比,综合实践活动课程具有主题开放、过程生成、成果多元等特征,这些特征在赋予课程活力的同时,也使“教、学、评”三者的协同运作面临更大挑战。

在实际教学中,综合实践活动课程的教学评割裂集中表现为三个层面:在“教”的层面,课程目标常以宏观素养表述,教师难以将其转化为每个教学环节可落地、可观测的学习任务,导致教学活动与课程目标之间缺乏清晰的对应关系;在“学”的层面,学生在开放性实践中缺少结构化引导工具,学习过程中的关键思维活动(如调研分析、方案推理、迭代归因等)往往“发生了但未留痕”,既无法被教师即时感知,也难以被学生自我觉察;在“评”的层面,由于过程性数据采集困难,评价普遍依赖终端成果展示和教师主观印象,既滞后于学习过程,也难以以为教学调整提供精准依据。<sup>②</sup>三者的断裂使得综合实践活动课程容易陷入“热热闹闹做活动、模模糊糊给成绩”的困境。

要打通教学评,需要解决两个关键问题:一是如何让学习活动本身同时成为评价证据的生成过程——即“学的过程就是评的依据”;二是如何让评价结果及时反馈,驱动教师调整教学和学生改进学习——即“评的结果服务于教和学”。人工智能技术和结构化工具的引入,为上述问题的解决提供了新的可能。本文以笔者在上海市徐汇中学实施的“AI 赋能·智慧校园”机器狗校园服务项目为案例,探索 AI 赋能下综合实践活动课程教学评一体化的实施策略,旨在提供可迁移、可推广的实践经验。

## 二、课程设计概要

### (一) 学情与背景

本课程于 2026 年春季在上海市徐汇中学实施。该校为上海市工程素养培育特色高中,学生信息技术素养整体较高。参与学生为高一年级 30 名选修生,约三分之一具有 Python 编程基础,其余学生无文本编程经验。

### (二) 项目设置

<sup>①</sup> 中华人民共和国教育部.教育部关于印发《中小学综合实践活动课程指导纲要》的通知[Z].教材〔2017〕4号,2017.该纲要明确要求综合实践活动课程从学生真实生活和发展需要出发,通过探究、服务、制作、体验等方式培养学生的价值体认、责任担当、问题解决和创意物化能力。

<sup>②</sup> 崔允漭,雷浩.教—学—评一致性三因素理论模型的建构[J].华东师范大学学报(教育科学版),2015,33(4):15-22.崔允漭指出,教学评一体化的核心在于目标统领下的教、学、评同向同行,评价应嵌入教学过程而非游离其外。

课程围绕三个真实的校园服务需求设置并行项目组（见表1），学生按兴趣自主选组，每组10人。编程采用Python语言“半成品代码模板”策略，学生只需在关键位置填入自主探究得出的方案参数，将技术门槛降至最低。

表1 三个项目组概况

项目组	项目名称	服务对象	解决的真实问题
后勤减负组	校园顺风狗	后勤人员	教辅资料经坡道搬运困难
校园文化组	赛博舞伴	文艺活动参与者	部分同学因怯场放弃上台表演
绿色校园组	暗夜巡更犬	值班教师、全校师生	晚自习后教室电器遗忘关闭

### （三）整体流程

课程共6课时（每课时45分钟），采用“唤醒（发现问题）→探究（调研分析）→转化（方案编码）→实践（测试迭代）→打磨（完善成果）→绽放（展示评价）”六阶段递进设计，每阶段均嵌入对应的过程性工具，使教学活动、学习任务和评价采集在同一时空同步发生。

## 三、AI 赋能教学评一体化的实施策略

### （一）以真实情境锚定教学评的共同起点

教学评一体化的前提是三者指向同一目标。<sup>③</sup>综合实践活动课程中，这个“共同目标”不应是抽象的素养表述，而应是一个学生可感知、可投入、可评判的真实问题。

本课程以“校园中谁需要帮助、需要什么帮助”为驱动问题开启教学。第一课时中，教师在机器狗现场演示后追问：“能不能教会它做一件对学校真正有用的事——不是好玩的事，是能帮到某个人、解决某个问题的事？”学生在讨论中自发提出了后勤搬运困难、同学怯场退缩、教室忘关灯等校园痛点，这些痛点经教师归纳后成为三个项目的起点。

这一设计的关键在于：真实问题同时锚定了教、学、评三者。对教师而言，“让机器狗完成一项有价值的校园服务”是教学活动的组织主线；对学生而言，围绕真实需求开展调研、设计和实施是自驱学习的动力源；对评价而言，“方案是否真正解决了问题”“服务对象是否真正受益”构成了可直接观测的终极评价标准。三者共同指向“在真实场景中解决真实问题”，天然对齐，无需人为“拉齐”。

实践中观察到，真实情境的锚定效应持续贯穿项目全程。例如，校园文化组在调研“怯场同学”的需求时，调研员在任务书中特别注明“不能太像采访，以免对方紧张，用语音备忘录做记录”——当服务对象是真实的人而非虚拟的情境时，学生的共情意识和伦理敏感性自然被唤醒。绿色校园组完成巡检方案后，学生主动找到学校保安教他们读懂程序输出——项目成果从课堂练习延伸为真实服务，同样源于“问题是真的、需求是真的”这一起点。

### （二）以结构化工具实现“学即评、评即学”

打通教学评的操作难点，在于找到一种载体，使学习活动和评价采集不再是两件分开的事，而是同一件事的两个面。本课程设计了四份结构化过程性工具，每份工具以“问题链”为核心骨架——学生沿问题链完成填写的过程即为深度学习活动，填写内容即为该阶段的评价证据（见表2）。

<sup>③</sup> 崔允漷.教—学—评一体化:意义与旨趣[N].中国教育报,2020-03-19(7).

表 2 过程性工具与教学阶段的对应关系

工具名称	嵌入阶段	学习支架功能	评价采集功能
《项目任务书》	唤醒	明确任务、角色分工	评估问题识别与责任意识
《实践探究引导单》	探究	问题链引导从数据到方案的转化	评估调研深度与方案设计质量
《测试记录表》	实践+打磨	记录“参数—表现—归因—修正”	评估迭代能力与证据推理能力
《小组互评表》	绽放	引导跨组学习与多元审视	采集同伴评价数据

以校园文化组的《测试记录表》为例，可以清晰看到“学即评”的运作机制：该组在第一轮测试中发现机器狗动作与音乐节拍不同步，在“原因分析”栏中写道“信号延迟，每次下达指令后有几秒间隔才执行动作”，随后在“修正方案”栏中记录了“代码前增加 `time.sleep(3)` 作为补偿”的解决策略。学生完成填写的过程，本身就是深度分析和迭代修正的学习过程；同时，教师批阅记录表时无需额外设计评价任务，直接可从“原因分析”栏判断学生是否具备分层归因能力——该组准确区分了“动作间断”（`sleep` 参数过大，属软件参数问题）与“节拍不同步”（信号延迟，属软硬件通讯问题）两个不同层面的原因，并分别采取了不同策略。这种分层归因的工程推理水平，在传统“凭印象给分”的评价方式下是很难被精确捕捉的。

这一机制使课程 60% 的评价权重有了切实的过程性数据支撑，每一份评价判断都可追溯至学生填写的原始记录。

### （三）以 AI 技术降低门槛、释放思维空间

综合实践活动课程教学评割裂的一个深层原因是技术门槛对思维活动的挤占——当学生将大量精力花在编程语法等技术操作上时，调研分析、方案设计、迭代反思等高阶思维活动的时间和空间就被压缩，教师也不得不将指导精力集中于技术问题而非思维引导。AI 技术的价值，正在于缓解这一矛盾。

本课程中 AI 赋能教学评一体化体现在两个层面：

第一，“半成品代码模板”降低编程门槛。教师为三个项目组分别准备了包含完整框架的代码模板，学生只需在 `# TODO` 标注处填入自己探究得出的参数。例如后勤减负组填入实测计算的坡度补偿角度 `lean_angle = 15`，绿色校园组填入讨论确定的光线阈值 `threshold = 150`。这一策略将编程简化为“填空”，使学生将精力集中于“填什么、为什么填这个数”——而后者恰恰是问题解决能力的核心。

第二，AI 大模型辅助代码生成。校园文化组在设计机器狗舞步时，借助 AI 大模型辅助生成了动作代码的初始框架。但 AI 生成的代码在真实场景中并不能直接使用——学生发现机器狗的实际响应存在约 3 秒的信号延迟，必须理解代码中每个参数的含义，才能完成针对性修正。真实场景的复杂性构成了防止 AI 依赖的天然屏障：AI 可以帮学生“起草”代码，但方案决策、参数校准和现场调试仍必须由学生自主完成。<sup>④</sup>

从教学评一体化的视角看，AI 技术的核心贡献在于：在“教”的层面，教师从重复性技术答疑中解放，将指导精力投向方案分析、迭代归因等思维引导；在“学”的层面，学生的认知资源从语法操作转向问题解决，过程性工具中“分析”“推理”“反思”类栏目的填写质量显著提升；在“评”的层面，过程数据采集更加完整，评价的精准度相应提高。三个层面的改善共同推动了教学评从割裂走向协同。

### （四）以迭代循环驱动深度学习与精准评价

<sup>④</sup> 余胜泉,胡韵.人工智能教师的发展阶段与路径[J].开放教育研究,2021,27(4):11-20.余胜泉等指出,AI在教育中应扮演“助手”而非“替代者”的角色,服务于学习者的深度学习。

传统综合实践活动课程中，“方案—实施—展示”的线性流程使学习在“第一版方案”完成后即进入展示环节，学生缺乏在真实环境中检验和修正方案的机会。本课程专门设计了两个课时用于实地测试与迭代优化，要求各组至少完成两轮“实测→归因→修正→复测”的完整循环。<sup>⑤</sup>

这一设计同时服务于“深度学习”和“精准评价”两个目标。

对深度学习而言，理论参数与物理执行之间的落差是最有力的学习触发点。三个项目组在首轮测试中均遭遇了方案与实际的显著偏差：后勤减负组机器狗上坡后仰（补偿角度不足且坡道入口有接缝冲击），校园文化组动作与节拍不同步（sleep 参数过大叠加信号延迟），绿色校园组出现检测误报（走廊环境光干扰导致阈值不当）。这些“失败”迫使学生直面理论与现实的差距，回到数据层面重新分析。校园文化组学生在记录中写下了一句颇具洞察力的话：“计算的结果与节拍并不意味着狗能达到结果动作与节拍。”<sup>⑥</sup>这句来自实践的领悟，标志着学生从“纸面正确”跨越到了“工程可行”的认知层次。

对精准评价而言，迭代循环使《测试记录表》成为最有深度的评价证据。教师通过对比同一小组前后两轮记录，即可直观判断：学生的归因是否从“笼统描述”进步为“分层分析”？修正方案是否与归因逻辑一致？第二轮是否真正解决了第一轮的问题？这种基于过程轨迹的纵向比较，比终端成果的横向评分更能反映学生的真实发展。

## 四、实践成效与反思

### （一）教学评一体化的实现程度

经过6课时的实施，教学评一体化在以下层面得到了较好实现：“学即评”层面，四份过程性工具为课程60%的过程性评价提供了可追溯的证据支撑，评价不再是教师的“印象分”。“评促教”层面，教师在批阅过程性工具时发现的共性问题，成为后续教学调整的依据。例如批阅引导单时发现后勤减负组对坡道入口接缝影响估计不足，遂在实测前提示该组重点关注过渡环节，实现了评价对教学的即时反馈。“评促学”层面，《测试记录表》中“原因分析”和“修正方案”两栏的结构化要求，迫使学生将模糊的直觉判断转化为有证据支撑的理性分析，填写过程本身即推动了思维深化。

### （二）AI 赋能的实际效果

AI 技术在课程中的作用主要体现在“释放”而非“替代”。<sup>⑦</sup>半成品代码模板和 AI 大模型辅助编程将学生从语法操作中解放，使其将认知资源集中于调研分析、方案设计和迭代反思等高阶活动。后勤减负组学生自主设计并 3D 打印了机器狗外挂载物装置，完成了教辅资料从 A 楼二楼到 B 楼一楼的跨楼搬运；校园文化组配合周杰伦《晴天》完成完整舞步表演，在全校科技节上获得广泛好评；绿色校园组完成高一年级 8 间教室用电巡检，异常检测准确率达 100%。三个项目组还不约而同地出现了超越课程预期的自主拓展行为：3D 打印硬件设计、引入 AI 大模型工具、主动教会保安使用检测程序——这些行为表明，当技术门槛不再是障碍时，学生的创造力和服务意识能够得到更充分的释放。

### （三）不足与改进

<sup>⑤</sup> 夏雪梅.项目化学习设计：学习素养视角下的国际本土化实践[M].北京:教育科学出版社,2018:86-95.夏雪梅强调,优质的项目式学习应包含完整的迭代优化环节,让学生在“做中学”“错中学”。

<sup>⑥</sup> 这一现象与 Grover 和 Pea 关于计算思维教育的主张相呼应:K-12 阶段编程教育的关键不在于语法教学,而在于培养学生的抽象建模与系统调试能力.参见: GROVER S,PEA R.Computational thinking in K-12:a review of the state of the field[J].Educational Researcher,2013,42(1):38-43.

<sup>⑦</sup> 杜建群.综合实践活动课程的核心素养培养功能与路径[J].课程·教材·教法,2018,38(4):48-53.杜建群指出,综合实践活动课程的核心能力需要在真实的活动中加以观察和评价,AI 辅助工具为过程性评价提供了新的技术支持。

本实践仍存在以下不足：第一，过程性工具以小组为单位填写，对组内不同角色学生个体贡献的区分度不足，后续可增设“个人贡献自述”板块。第二，AI 工具使用需要更明确的引导边界。真实场景的复杂性虽能在一定程度上防止 AI 依赖，但教师仍需在课程中引导学生区分“AI 可以帮你做什么”和“你需要自己想清楚什么”。第三，本实践以质性分析为主，缺乏前后测量化对比，后续可引入标准化问卷进一步检验课程效果。第四，研究仅在一个班级 30 名学生中实施，结论的推广性有待扩大样本后验证。

## 五、结语

综合实践活动课程的教学评一体化，核心不在于额外增加评价环节，而在于让学习活动本身成为评价证据的自然生成过程。本文基于高中机器狗校园服务项目的实践，提炼出以下四条可迁移的策略经验：

一是用真实问题锚定教学评的共同起点。当教学目标、学习任务和评价标准共同指向“在真实场景中解决真实问题”时，三者天然对齐，无需人为“拉齐”。

二是用结构化工具实现“学即评”。以问题链为骨架的过程性工具，使学生的填写过程同时完成学习推进和评价证据生成，做到一次活动、双重功能。

三是用 AI 技术释放师生的认知资源。AI 的价值不在于替代思维，而在于将师生从重复性技术操作中解放，使教师聚焦思维引导、学生聚焦深度探究，从而提升教学评各环节的质量。

四是用迭代循环让“失败”成为教学评的共同资源。测试中的“失败”既是驱动深度学习的触发点，也是生成高质量评价证据的关键环节，课程设计应有意识地为迭代留出空间。

以上策略并不局限于机器狗项目或人工智能课程，任何以真实问题为载体的综合实践活动课程，均可参考这一“真实情境—结构化工具—AI 辅助—迭代循环”的框架，探索适合自身条件的教学评一体化实施路径。