

WIP: 通过 AI 驱动的同伴代理增强基于游戏的学习

Chengzhang Zhu¹, Cecile H. Sam², Yanlai Wu¹, Ying Tang^{1†}

¹Department of Electrical and Computer Engineering, Rowan University, USA

²Department of Educational Leadership, Rowan University, USA

Email: zhuche95@students.rowan.edu, sam@rowan.edu, wuyanl37@rowan.edu, tang@rowan.edu

摘要—这篇进展中的论文介绍了 SPARC (系统性问题解决和算法推理儿童版), 这是一个旨在提高 K-12 STEM 教育参与度与知识保留的游戏化学习平台。传统方法常常难以激发学生兴趣或促进深层次理解, 尤其是对于复杂的科学概念。SPARC 通过将互动、叙事驱动的游戏与基于大型语言模型的人工智能同伴代理相结合来解决这些挑战。该代理不仅仅是提供答案, 而是会与学生进行对话和探究, 促使他们解释概念并共同解决问题。平台设计基于教育理论, 并且与州立学习标准紧密相连。初步的课堂试点采用了结合前测和后测、游戏内分析以及来自学生和教师的定性反馈的多方法评估框架。初步结果显示, SPARC 显著提高了学生的参与度, 大多数参与者表示对 STEM 科目的兴趣增加, 在后测结果中观察到概念理解程度有所提高。正在进行的发展工作集中在完善人工智能代理、扩展课程整合和提升可访问性上。这些早期成果展示了将基于 AI 的同伴支持与游戏化学习相结合, 为 K-12 学习者创造包容、有效且吸引人的教育体验的巨大潜力。

Index Terms—K-12 教育, 学生参与, 基于游戏的学习。

I. 介绍

这篇研究到实践的论文描述了 SPARC (系统性问题解决和算法推理儿童平台) 的实施及其初步评估, 这是一个创新的人工智能驱动的游戏化学习平台, 旨在通过沉浸式互动体验来提升 K-12 阶段的 STEM 教育。

传统的 K-12 STEM 教育经常面临重大挑战。一个主要的挑战是保持学生的参与度, 因为传统教学方法通常让学生处于被动学习的角色, 限制了他们的动机和兴趣。此外, 由于场景的局限性, 传统课堂方法往往难以确保有效的知识转移 [1], 导致学生无法将其理论知识应用于实际的真实世界场景。

This work has been accepted for presentation at IEEE Frontiers in Education Conference 2025. The final published version will be available via IEEE Xplore.

为了弥合这些教育差距, SPARC 项目引入了一个以叙述为中心的交互式学习环境。SPARC 当前研究的核心是人工智能驱动的同伴代理。它们不仅作为被动的导师 [2] 进行叙事指导, 而是作为积极的同伴参与其中。在我们的设计中, 它们可以通过结构化的对话和基于探究的课程 [3] 提示并引导学生。这种将 AI 同伴代理的独特整合旨在通过将被动学习转变为动态的学生主导探究来增加学生的参与度、批判性思维和概念理解。

在以下各节中, 我们概述了指导这一实践的理论基础, 详细描述了我们的实施过程, 讨论了初步发现, 并强调了这种创新教育方法的影响和未来方向。

II. 背景与动机

教育研究越来越支持使用互动和个性化的方法来解决传统限制。利用好奇心、挑战性和乐趣等内在动机的博弈化学习环境已经显示出显著的潜力, 可以增强学生的学习动力和参与度, 相比传统方法 [4] [5]。具体来说, 游戏促进积极的学习体验, 提供即时反馈, 鼓励解决问题, 并促进成就感——所有这些因素都与改善学习成果相关联 [6] [7]。

此外, 最近的理论发展强调了同伴学习和社会互动在教育环境中的重要性。根据联通主义理论, 学生在通过与同龄人 [8] 进行协作沟通和交互资源主动建立概念之间的联系时, 学习效果最佳。这样的互动鼓励学生表达他们的理解, 并促进知识的更深层次理解和保留 [9]。

人工智能 (AI), 如大型语言模型, 为个性化教育增添了另一个重要的维度。由 AI 驱动的系统可以根据每个学习者的个人需求自适应地定制内容和反馈, 在大规模范围内有效提供个性化的学习体验 [10] [11]。如果将由 AI 驱动的同伴代理集成到游戏环境中, 它们可以

提供一个独特的主动对话和个人支持机会 [12], 结合了游戏的动机性益处与 AI 的自适应性益处。

这些洞察力推动了 SPARC 项目, 旨在利用 AI 驱动的同伴代理在游戏化的 STEM 学习体验中, 特别针对服务不足的学生群体的参与度和概念理解。这一正在进行的工作探讨了这种创新方法的早期实施及其初步效果, 以解决传统 K-12 STEM 教育中的缺口。

III. 研究到实践框架以及相关方面

SPARC 的设计基于已建立的学习理论, 这些理论强调积极的、情境化的和社会中介的学习。从建构主义的角度来看, 学生通过动手实践和反思来构建意义的方式学习最有效, 而不是被动地接收信息 [13]。根据这一原则, SPARC 整合了边做边学的活动和基于问题的挑战, 以在有意义的情境中促进概念理解。

此外, 社会学习理论 [14], 特别是那些强调同伴合作和情境互动的理论, 通过人工智能驱动的同伴代理和协作游戏任务等功能影响游戏的设计。这些元素促进了表达、建模和相互支持——这些都是在协作环境中实现更高层次学习的关键机制。

认知学习理论也通过告知系统的反馈使用、记忆强化和元认知提示 [15] 发挥着至关重要的作用。例如, SPARC 的同伴代理积极促使学生解释他们的推理过程, 这一做法已被证明可以增强记忆保留和概念清晰度。

这些互补的理论基础共同指导了 SPARC 的核心机制, 确保游戏不仅具有吸引力, 还与经验支持的教学策略相一致。

A. 设计原则在 SPARC 中的应用

在此基础上, SPARC 应用教育设计的核心原则来创建既吸引人又具有教学效果的活动。该平台的内容完全符合新泽西州学生学习科学标准 (NJSL-S)。

1) 沉浸式叙事情境: SPARC 采用以故事驱动的方法来构建学习框架。受《神奇校车》的启发, 游戏的故事线将玩家缩小到乘坐宇宙飞船探索人体的程度, 在此过程中与细胞和器官互动, 如图 1 所示。通过在冒险中嵌入课程, SPARC 利用讲故事的方式吸引注意力, 并为指导旅程的任务赋予意义。

尽管 SPARC 的故事线包括探索人体这样的创意元素, 如同乘坐宇宙飞船一般, 所有教育内容、对话和环境都经过生物学家专家的全面审核。这确保了能够建立

引人入胜的真实世界联系, 同时不会牺牲科学准确性或引入误解。

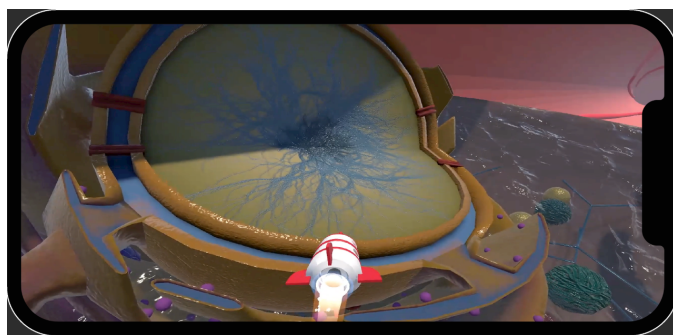


图 1. 火箭探索。

2) 人工智能同行代理交互: 该游戏特色在于拥有一位智能同伴代理, 作为对话伙伴, 而不是传统的非玩家角色 (NPC)。这位由人工智能驱动的角色陪伴学生在游戏中穿行, 引导故事情节和玩家的探索。随着游戏的进展, 该代理通过在适当的时候提示和提问来建立与玩家的信任并强化玩家的记忆, 如图 2 所示。



图 2. 代理交互。

3) 协作与社会学习: SPARC 支持团队合作和个人探索, 以反映现实生活中的学习环境。该游戏允许学生组成队伍 (最多四人) 一起解决问题、分享线索并查看集体和个人进度。内置聊天和团队目标鼓励玩家在游戏过程中进行同伴讨论与协作, 培养沟通技巧和同伴学习。同时, 独立解谜任务支持独立学习, 确保合作不会干扰个人问题解决练习。这种平衡有助于适应不同的学习偏好, 并将社会互动作为学习工具加以推广。

4) 游戏玩法多样性与参与系统: 为进一步鼓励学生的持续参与, SPARC 整合了多种引人入胜且可重复游玩的游戏模块。例如, 心血管模块包括一个基于节奏的音乐游戏, 而氧气运输活动则被设计为一项高速赛车

挑战。这些互动元素旨在使学习变得有趣，并激励学生反复游玩。此外，SPARC 还拥有成就系统，该系统会根据学生不同玩法中达到的里程碑奖励他们一些外观解锁项目，同时还有全球排行榜来促进玩家之间的友好竞争。平台持续更新新内容和用户驱动改进，确保学生有持续的理由重新参与并探索新的挑战。这些机制将 SPARC 从一次性体验转变为一个动态且持久的学习环境。

B. 迭代设计与课程对齐

将研究转化为实践的一个重要方面是迭代的以用户为中心的发展过程。SPARC 团队与 K-12 教师和学生多个设计周期中进行了合作。持续的教师反馈确保了游戏内容紧密符合 NJSLs，导致有针对性的教学调整，例如在血液循环之旅中增加了力与运动的内容。每个 SPARC 模块都是特意设计来支持基于标准的教学，当前的内容集中在循环系统上，新的模块（如消化系统）正在积极开发中。同时，在测试过程中学生的反馈带来了游戏机制和界面可用性的实用改进，比如可定制的视角和重置功能，这些解决了早期原型中的晕动症问题。

C. 同伴代理在实践中的应用——通过教学学习

人工智能同伴代理（“CAP”）是 SPARC 最具创新性的特性之一，将同伴辅导和可教学代理的教育研究转化为引人入胜的游戏机制。CAPs 不是进行讲授，而是作为同龄人与学生互动——提问、寻求澄清并促使学习者解释概念，通过让学生扮演教师的角色来深化理解。这一方法基于证据表明教他人可以增强学习效果，使课堂互动既具有激励性又富有教学意义。老师和学生报告称，CAPs 创造了更加沉浸式和教育性的游戏环境。

技术上，CAP 代理是基于本地部署的经过教育用途微调的 Llama3 语言模型实现的一个模块化系统。所有学生行为，包括导航、任务完成和对话，都带有元数据记录下来，以支持学习分析和代理的内部记忆模块。该记忆追踪短期和长期上下文，根据时间、重要性和与每个学生不断演变的知识地图的相关性对事件进行索引，并计算好奇心指数以实现自适应提示。模块化设计还允许未来集成基于人类反馈的强化学习和情绪识别，进一步增强代理提供个性化支持的能力。

IV. 评估方法

为了评估 SPARC 的教育影响和实际实施，我们采用了一个结合了学习成果、与 AI 驱动的同伴代理互动以及学生参与度的定量和定性衡量方法的多方法评估框架 [16]。

A. 预测试和后测试评估

学生的学习进步是通过游戏前和游戏期间进行的预测试和后测试来衡量的。这些评估由教师提供信息并经过审查。评估涵盖了关键的 STEM 主题，并允许对概念理解进行直接比较。结果被分析以确定可归因于 SPARC 的知识改进。

B. 游戏中交互分析

游戏中的分析捕获了模块完成率、互动频率、任务时间以及参与模式等指标。这些行为数据为持续优化游戏机制和适应性反馈提供了信息。目前，数据分析是事后进行的，但未来的迭代将整合实时分析以提供适应性的游戏支持并实现更深层次的代理集成。

C. 学生代理对话分析

对学生代理对话进行的定性分析侧重于概念准确性、解释深度以及词汇的有效使用。记录被系统地编码，以评估学生与同伴代理的互动如何强化 STEM 学习并解决误解。

D. 学生反馈调查

调查收集了学生关于学习感知、动机、可用性和课堂整合的描述性反馈。对定性响应进行主题分析有助于识别反复出现的优势、改进领域和实施的实际考虑。

为了补充学生调查，游戏结束后进行了 4 个焦点小组（每个小组人数在 3-6 人之间），以更全面地了解他们的体验。学生们讨论了与游戏内容相关的首选学习方式、改进建议以及该游戏在其整体电子游戏体验中的位置。

E. 学生焦点小组

通过整合这些评估方法，并利用田野笔记和观察进行三角互证 [17]，我们获得了对 SPARC 有效性的一个全面理解，并持续为后续发展提供信息。

V. 初步发现

尽管 SPARC 项目仍在进行中，多次课堂试点已为其教育潜力提供了宝贵的见解。具体而言，2024 年和 2025 年的试点实施允许在主观反馈和客观绩效指标方面分阶段收集数据。

A. 学生对 2024 年和 2025 年试点的看法

两轮 SPARC 部署通过学生调查收集了感知数据。2024 年的首次试点涉及 56 名学生，而 2025 年的后续试点从 54 名学生那里收集了反馈。这些调查重点关注学生报告的学习效果、乐趣、感知的有用性以及与传统教学的比较。

在两个试点项目中，结果表明学生参与度高，并且对平台有积极的看法。在 2024 年早期版本的试点中，超过 80% 的学生表示对 STEM 领域的兴趣增加，89% 的人认为 SPARC 有效或优于传统教学。大约 72% 的人相信他们会向朋友推荐这款游戏。

在表 I 中显示的 2025 版本最近的一项调查中，共有 54 名学生（29 人自认为男生，25 人自认为女生）报告了从中学习到了一定程度的知识（M = 3.81，在五分制下）。关于解决问题的能力，有所提高是最常被报告的回答（51.9%），其次是大大提高（20.4%）。大多数学生对他们的整体游戏体验描述为正面的，其中 42.3% 的人认为很有趣，30.8% 的人认为中立，没有人觉得不有趣。此外，大多数人表示他们相信这款游戏可以帮助同龄人学习这些主题，并且许多人表达了在玩游戏后希望了解更多关于人体工作原理的兴趣。

表 I
学生对 SPARC 游戏体验的感知总结

Survey Item	1	2	3	4	5
Overall learning from SPARC	0.0%	1.9%	33.3%	46.3%	18.5%
Improved problem-solving skills	5.6%	14.8%	51.9%	20.4%	7.4%
Overall enjoyment of the gameplay experience	0.0%	3.8%	30.8%	42.3%	23.1%
SPARC compared to traditional assignments	1.9%	3.7	33.3%	44.4%	16.7%

注意：5 点李克特量表的响应选项。对于项目 1-3，1=根本不，5=相当多。对于项目 4，1=远不如传统作业，5=好得多。N = 54。

B. 学习成果与认知收获

除了调查数据外，2025 年的试点还包括前测和后测评估以评价概念理解情况。共有 60 名学生在游戏开始前完成了前测，其中 44 人在与 SPARC 模块互动后提供了有效的后测数据。

模块的主要差异在于交互风格和教学方法，特别是学生是与采用自适应 LLM 生成查询提示的 AI 同伴代理互动（叙事故事模块），观看被动的教学视频（被动视频教学模块），还是直接参与动手游戏任务（交互式游戏模块）。表 II 总结了这些类型的准确度得分。

表 II
模块类型（2025）的预测试和后测试准确性（%）

Module Type	Pre-test (%)	Post-test (%)
Agent-based Module(LLM Inquiry)	39.2	45.7
Agent-based Module(No LLM Inquiry)	39.2	24.4
Passive Video Instruction Modules	50.0	28.3
Interactive Gameplay Modules	47.3	59.2
Overall	46.6	49.7

注意：前测 N=60；后测 N=44。

值得注意的是，完成代理支持（LLM Inquiry）模块的学生在后测试中的得分与其它组相比有所提高，尽管这些模块的概念难度更大。与其他模块相比，LLM Inquiry 模块还包含更深入且认知需求更高的问题，这可能是导致观察到的进步的原因之一。相比之下，仅使用视频的模块显示出较低的后测试分数，表明被动格式对于学习新概念的效果较差。我们还注意到一些前测分数高于预期，因为学生讨论了答案或在网上查找了答案。这可能有助于解释后测试分数下降的情况，特别是对基于视频的模块而言。

综合来看，这些研究结果表明，交互式、代理驱动的学习环境比非代理或基于视频的模块更能支持学生理解复杂主题。未来的试点将进一步加强实验控制，并纳入直接对照组以更严格地评估归因于 SPARC 的学习增益。

C. 学生积极的反馈

对学生在开放性问题 and 焦点小组中定性回应的分析显示，他们对 SPARC 的互动学习体验持压倒性的积极态度。与调查结果一致，学生报告称，SPARC 的经历是一种新颖的学习体验，他们认为这是一种很好的补

充教师教学的方式。在焦点小组中，学生们希望有更多的像 SPARC 这样的互动活动，但也重视老师提供基础信息的作用。学生们表达了高度参与的态度，表示，“游戏让我保持专注，并帮助我更有效地学习”，“我发现探索和理解人体的不同部分很令人兴奋。”

通过 SPARC 的初始实现，已经识别出若干实际和技术挑战，为未来的改进提供了关键教训。

VI. 挑战与经验教训

A. 教育实施挑战

课堂实施显示了显著的挑战：

- 有限的设备可用性和不一致的技术基础设施（例如不可靠的 Wi-Fi 和共享设备）阻碍了稳定的实施。
- 技术难题导致数据收集不完整，影响了评估的准确性和完整性。

为了解决这些问题，未来的 SPARC 版本将包括一个离线模式，使学生可以在没有持续互联网连接的情况下访问核心学习模块。该平台还在优化以实现与更多设备的兼容性，包括平板电脑、智能手机和共享课堂设备，以在各种教育环境中最大限度地提高可访问性。此外，还将为教师提供一份技术检查表和故障排除指南，以便快速诊断和解决常见的连接或硬件问题。这些改进旨在最大程度减少技术干扰，并确保所有学生都能获得公平一致的访问。

首先，初始用户界面（UI）设计对某些学生来说存在易用性困难，导致了混淆和导航挑战。其次，虽然总体上有效，但由 AI 驱动的同伴代理在实时响应能力、准确性和交互自然度方面仍需进一步改进。

B. 技术与用户体验挑战

在这种情况下，未来的研究与发展将重点放在完善用户界面交互和提升 AI 模型上，旨在提供更加直观和无缝的学生体验。

VII. 未来工作与影响

展望未来，我们将专注于扩大评估范围并完善 SPARC 系统以应对这些已识别的挑战。计划中的下一步包括在更多的教室和学校进行更大规模的试验，以收集可靠的数据，验证初步的积极趋势，并更自信地评估知识的增长和保持情况。通过增加样本量和研究时长，我们可以更好地评估学习成效并对平台进行相应的微调。

SPARC 的人工智能驱动特性，尤其是同伴代理功能，将进一步增强，以提供更加个性化和适应性的支持给不同类型的学员。我们正在开发一个独立的、模块化的代理模块，可以灵活地集成到各种教育游戏和平台中，使其能够适应任何学科领域或数字环境。这种模块化设计旨在最大限度地提高人工智能支持下的同伴学习在教育领域的广泛应用性和影响力。

另一个重点是深化课程整合。我们继续与教育工作者合作，使 SPARC 内容与学校标准保持一致，并提供指南和培训以实现无缝课堂整合 [18]。通过提高可访问性、AI 适应性和课程匹配度，SPARC 致力于将其影响扩展到初始试点之外。

最终，像 SPARC 这样的游戏化人工智能平台，特别是通过可访问的移动技术，有可能为 K-12 学习者普及 STEM 教育。我们预计从 SPARC 中学到的经验将有助于指导未来人工智能教育工具的设计，并激发下一代科学家和工程师。

VIII. 结论

本在研论文介绍了 SPARC 的初步实现和评估，这是一个通过 AI 驱动的同伴代理增强的游戏化学习平台。早期结果表明，SPARC 可以通过互动对话和沉浸式游戏显著提升学生对 STEM 的兴趣和参与度。尽管实际和技术挑战仍然存在，但该平台展示了利用 AI 增强游戏化推进 K-12 教育的强大潜力。持续的发展和更大规模的研究将进一步完善 SPARC，并支持其在包容性 STEM 学习中的更广泛影响。

致谢

这项工作部分得到了 Google 学术研究奖和罗文大学-罗格斯-凯姆登州长委员会的支持。

参考文献

- [1] J. Hu, “The challenge of traditional teaching approach: A study on the path to improve classroom teaching effectiveness based on secondary school students’ psychology,” *Lecture Notes in Education Psychology and Public Media*, vol. 50, pp. 213–219, 04 2024.
- [2] R. Hare, Y. Tang, and C. Zhu, “Combining gamification and intelligent tutoring systems for engineering education,” in *2023 IEEE Frontiers in Education Conference (FIE)*, 2023, pp. 1–5.
- [3] M. Fitzmaurice, “Considering teaching in higher education as a practice,” *Teaching in Higher Education*, vol. 15, no. 1, pp. 45–55, 2010. [Online]. Available: <https://doi.org/10.1080/13562510903487941>

- [4] T. Laine and R. Lindberg, "Designing engaging games for education: A systematic literature review on game motivators and design principles," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 804–821, Oct. 2020.
- [5] C. Franzwa, Y. Tang, A. Johnson, and T. Bielefeldt, "Balancing fun and learning in a serious game design," *International Journal of Game-Based Learning*, vol. 4, pp. 37–57, 10 2014.
- [6] N. Peña-Miguel and M. Sedano Hoyuelos, "Educational games for learning," *Universal Journal of Educational Research*, vol. 2, no. 3, pp. 230–238, 2014.
- [7] B. Paras and J. Bizzocchi, "Game, motivation, and effective learning: An integrated model for educational game design." 01 2005.
- [8] K. M. Dunaway, "Connectivism," *Reference Services Review*, vol. 39, no. 4, pp. 675–685, 2011. [Online]. Available: <https://doi.org/10.1108/00907321111186686>
- [9] S. Papadakis, A. Trampas, K. Barianos, M. Kalogiannakis, and N. Vidakis, "Evaluating the learning process: The "thimeledu" educational game case study," 01 2020, pp. 290–298.
- [10] T. Kabudi, I. Pappas, and D. H. Olsen, "Ai-enabled adaptive learning systems: A systematic mapping of the literature," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100017, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666920X21000114>
- [11] J. Liang, R. Hare, T. Chang, F. Xu, Y. Tang, F. Wang, S. Peng, and M. Lei, "Student modeling and analysis in adaptive instructional systems," *IEEE Access*, vol. 10, pp. 59 359–59 372, 2022, publisher Copyright: © 2013 IEEE.
- [12] J. Wang, Y. Tang, R. Hare, and F.-Y. Wang, "Parallel intelligent education with chatgpt," *Frontiers of Information Technology & Electronic Engineering*, vol. 25, no. 1, pp. 12–18, 2024. [Online]. Available: <https://doi.org/10.1631/FITEE.2300166>
- [13] P. A. Ertmer and T. J. Newby, "Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective," *Performance Improvement Quarterly*, vol. 26, no. 2, pp. 43–71, 2013. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/piq.21143>
- [14] M. Zhou and D. Brown, *Educational Learning Theories: 2nd Edition*. GALILEO Open Learning Materials, 2015, education Open Textbooks, Book 1. [Online]. Available: <https://oer.galileo.usg.edu/education-textbooks/1>
- [15] J. Bransford, N. Vye, R. Stevens, P. Kuhl, D. Schwartz, P. Bell, A. Meltzoff, B. Barron, R. Pea, B. Reeves, J. Roschelle, and N. Sabelli, "Learning theories and education: Toward a decade of synergy," 01 2005.
- [16] M. Q. Patton, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, 4th ed. Thousand Oaks, CA: SAGE Publications, 2015. [Online]. Available: <https://collegepublishing.sagepub.com/products/qualitative-research-evaluation-methods-4-232962>
- [17] J. Phillippi and J. Lauderdale, "A guide to field notes for qualitative research: Context and conversation," *Qualitative Health Research*, vol. 28, no. 3, pp. 381–388, 2018.
- [18] L. Cui, C. Zhu, R. Hare, and Y. Tang, "Metaedu: a new framework for future education," *Discover Artificial Intelligence*, vol. 3, no. 1, p. 10, March 2023. [Online]. Available: <https://doi.org/10.1007/s44163-023-00053-9>