

大模型驱动的智能系统对小学科学课堂活动导学的影响研究

A Study on the Impact of LLM-Driven Intelligent Systems on Primary School Science

Classroom Design

吴涵¹, 欧阳雨彤¹, 张立山^{1*}

¹ 北京理工大学教育学院

*lishan.zhang@bit.edu.cn

【摘要】 随着大语言模型在教育领域的广泛应用, 如何利用其智能化特质调节并提升课堂活动导学成效成为研究热点。本研究开发了一套面向小学科学课《水结冰了》的大模型驱动的智能系统, 通过对北京某小学四年级两个班级开展准实验研究, 结合系统日志、对话记录及课堂视频等进行数据分析。结果表明, 使用智能系统的实验班在保证探究深度的前提下, 课堂活动导学效率显著提升, 教学耗时较对照班减少了 15%, 并且系统内置的智能反馈帮助 46.2% 的小组有效纠正了实验观察错误, 显著优于和通用 AI 对话的对照班。本研究验证了大语言模型驱动的智能系统在提升课堂活动导学与学生科学探究质量方面的协同价值, 为生成式人工智能进入常态化课堂提供了实践参考。

【关键词】 课堂活动导学; 大语言模型; 小学科学探究; 智能系统

Abstract: With the widespread application of large language models in education, how to leverage their intelligent features to adjust classroom arrangements and improve the effectiveness of activity-based learning has become a research hotspot. This study developed a large-model-driven intelligent system for the elementary school science lesson "Water Freezes." A quasi-experimental study was conducted with two fourth-grade classes in a Beijing elementary school, and data analysis was performed using system logs, dialogue records, and classroom videos. The results showed that the experimental class using the intelligent system significantly improved the efficiency of activity-based learning while maintaining the depth of inquiry. Teaching time was reduced by 15% compared to the control class. Furthermore, the system's built-in intelligent feedback helped 46.2% of the groups effectively correct experimental observation errors, significantly outperforming the control class that interacted with general AI. This study verifies the synergistic value of large language model-driven intelligent systems in improving the quality of activity-based learning and students' scientific inquiry, providing a practical reference for the integration of generative artificial intelligence into regular classrooms.

Keywords: classroom orchestration, LLM, elementary school science inquiry, intelligent systems

1. 引言

在小学科学教育中, 探究式学习是培养学生科学素养的核心路径(任长松, 2005)。然而, 在《水结冰了》等涉及复杂现象观察与实时数据记录的科学实验课中, 教师通常会面临“课堂活动导学困境”: 一方面难以实时监控多个小组的探究进度与认知误区; 另一方面, 传统的预

设支架难以应对学生在实验过程中生成的个性化数据。虽然通用大语言模型展示了强大的交互潜力,但由于缺乏科学情境的约束,其在课堂活动导学中的精准度与系统性仍有待提升。

课堂活动导学强调教师在多约束情境下对教学活动的实时管理(Dillenbourg & Jermann, 2010)。研究认为,大语言模型(large language model, LLM)不应仅作为学生的“对话助手”,更应成为教师的“感知增强器”,帮助教师调节课堂活动的安排并辅助其进行教学决策。为此,本研究设计并实现了一个大模型驱动的智能系统,其核心创新在于:通过结构化任务流驱动,将 LLM 的推理能力嵌入小学科学课堂中的四个关键环节,并利用教师端的 AI 总结分析功能将学生数据转化为可行动的教学洞察。通过对比使用该智能系统的实验班与使用通用大模型“豆包”的对照班的表现,本研究试图揭示大语言模型驱动的智能系统在课堂活动导学以及学生实验探究表现方面的影响作用,为未来智能化课堂的设计提供理论支持与实践依据。

2. 相关研究

2.1. 课堂活动导学系统

课堂活动导学(Classroom Orchestration)是指教师在充满多约束的课堂环境中,对多层次教学活动进行实时管理与监控的复杂过程。从宏观抽象角度来看,活动导学本质上是一种高度依赖情境感知且具有动态灵活性的适应性教学调节机制(Dillenbourg, 2013)。为了支持教师在复杂的社交层面对课堂进行实时监控(VanLehn et al., 2021),课堂活动导学系统(Classroom Orchestration system, COS)应运而生。现有研究表明, COS 在辅助教师管理大规模授课情境、提升教学稳健性以及优化学习成效方面具有显著价值(Feng, Zhang, Wang, & Cai, 2023)。然而,传统的活动导学系统往往依赖于预设的规则引擎,难以对学生生成的非结构化探究内容进行深度语义解析。

基于此,本研究设计的课堂活动导学系统以大语言模型为核心驱动,旨在将智能推理能力和实时数据分析能力嵌入教师的课堂中。系统不仅遵循课堂授课的原始流程进行模块化导学,更通过对学生学习进度与学习成效数据的后台汇聚,实现了高度透明化的学情监控。这种设计能够显著增强教师对课堂动态的感知维度,从而在多约束的科学探究情境下,更有效地达成教学任务导学与课堂活动治理的目标。

2.2. 大语言模型赋能的科学探究支架研究

在科学探究教育中,提供及时的教学支架是促进学生概念转变和教师指导教学过程的关键(Van Uum, Verhoeff, & Peeters, 2017)。随着生成式人工智能的发展,大语言模型凭借其深层的语义理解与情境生成能力,为个性化课堂导学活动的设计提供了新的可能性(Sharma, Mittal, Kumar, & Bhardwaj, 2025)。现有研究多关注 LLM 作为通用对话助手在知识问答中的表现,但在特定学科情境下,如何利用 LLM 对学生生成的实验观测数据进行实时反馈和促进教师课堂调节仍处于探索阶段。尤其是在小学科学课堂中,学生的语言表达往往具有模糊性与碎片化特征, LLM 的引入能够有效监控学生的学习状态,并基于实时数据提供具有针对性的引导。此外,目前的研究较少关注 LLM 如何与具体的课堂活动导学流程深度耦合。本研究通过构建特定领域的提示词策略,使 LLM 能够精准识别科学课堂中的关键数据特征,从而实现从“通用对话”向“学科专业支架”的转化。

3. 系统设计

3.1. 技术方案

智能系统的设计采用云端驱动的服务架构,其核心建立在底层基础模型与上层教学应用的交互基础之上。在整体架构设计上,系统构建了由模型支撑层、业务逻辑层与前端应用层

组成的协同体系。系统底层接入通义千问大语言模型，利用其卓越的语义理解和逻辑推理能力，为科学探究中的复杂情境提供智能化处理支撑。通过 API 调用机制，系统能够实时解析学生在实验过程中生成的数据，并根据预设的相关教学资料生成针对性的引导建议。

业务逻辑层作为系统的中枢，通过提示词工程（prompt engineering）将特定的学科知识、教学目标与实时学生数据进行深度融合，确保大模型的输出能够精准契合小学科学课的教学范式。同时，系统利用数据中转机制，实现了数据的结构化存储与处理，为后续的教学总结与数据可视化提供了数据基础。

为了满足课堂活动导学的实际需求，系统在应用层划分为相互关联的学生端与教师端。学生端系统依托移动学习终端，侧重于实验过程的引导与自主探究，将大模型生成的反馈即时作用于学生的探究行为；教师端系统则以电脑为载体，通过聚合大模型的总结能力与数据可视化技术，将各小组的探究状态凝练为直观的教学洞察。这种双端协同的技术架构不仅增强了学生的自我调节学习能力，更为教师实时感知课堂动态、进行课堂活动导学的调节提供了稳固的技术底座。

3.2. 功能架构

本系统紧扣科学探究的典型环节，将课堂教学流程转化为数字化任务流，构建了由“材料选择、实验模拟、水温记录、观察解释”组成的四个活动模块。

3.2.1. 学生端

学生端以“任务驱动”为设计核心，通过交互式界面引导小组协作，其具体功能如下：

(1) 材料选择：如图 1 所示，系统提供三种可选的制冷材料及相应的科学背景资料。学生需在阅读资料后，在线提交选定材料的理由及初步实验设计方案。提交后，系统通过 LLM 语义分析对学生的提交结果进行实时评测，并提供一个智能反馈。该反馈不直接提供标准答案，而是指出学生提交数据中的问题并引导其自我修正。模块支持多次提交迭代，实现了小学生基于 AI 反馈的自我调节学习过程。

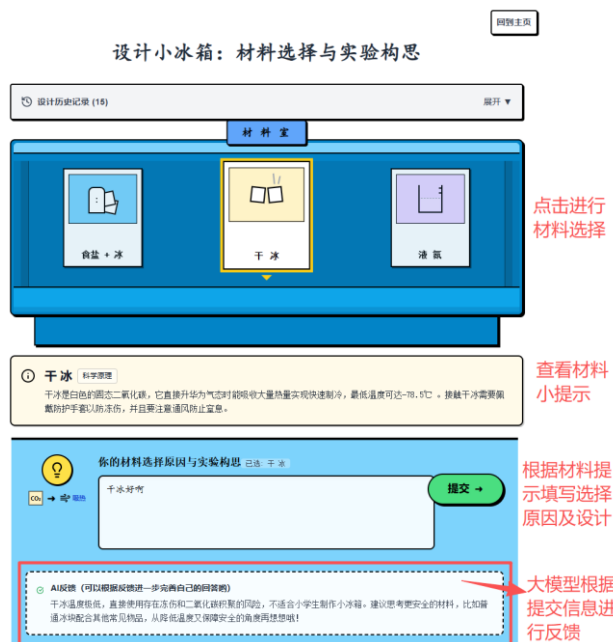


图 1 学生端材料选择模块

(2) 实验模拟：为了降低学生进入正式实验前的操作错误率并启发学生思考，该模块构建了一个虚拟仿真环境，如图 2 所示。系统提供大烧杯、试管、铁架台及温度计等数字化器材，并内嵌了温度计读书等关键操作提示。学生可以通过交互式拖拽来模拟实验装置的搭建

及操作流程。该板块的模拟可以进行实验重置，通过数字化的预演，来帮助学生规范正式实验操作的步骤。

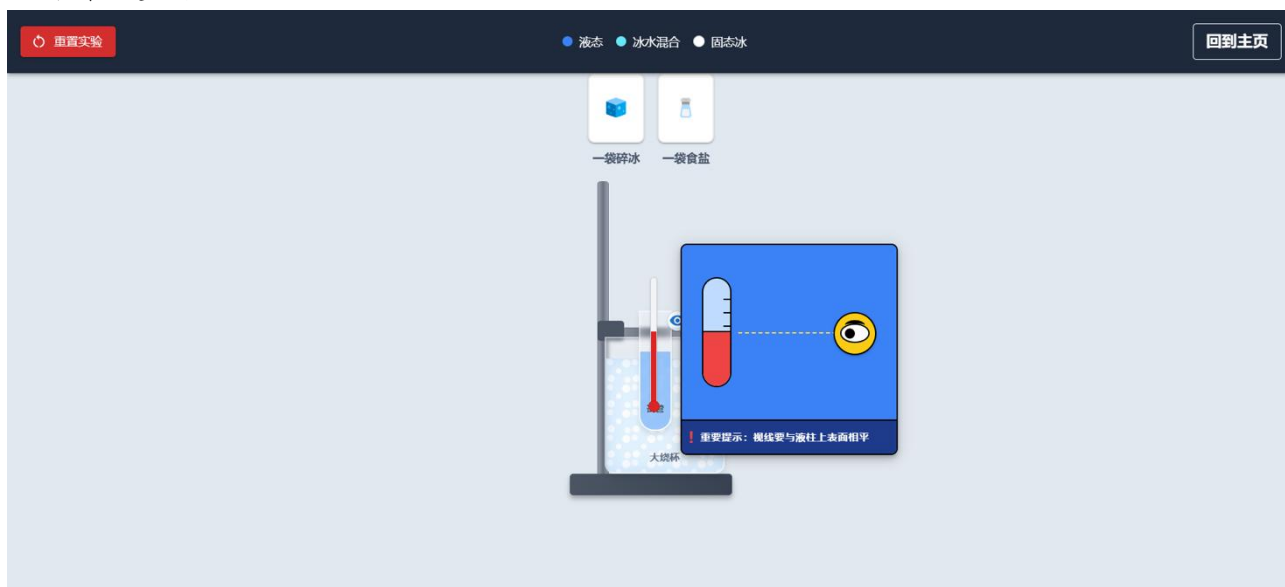


图2 学生端实验模拟模块

(3) 水温记录：该模块对应正式的科学实验环节，系统提供了结构化的数据录入接口，如图3所示。学生需同步观测实验现象，并周期性记录从0至7分钟内关键节点的实时水温及物理形态变化。该模块还可以将学生输入的离散数据实时同步至教师端，为后续的数据可视化与AI总结提供支撑。

▲ 水温记录

[回到主页](#)

观察并记录水结冰过程中试管内水温的变化和出现的现象。

时间	水温	现象	确认
第0分钟 (min)	20 摄氏度 (°C)		✓
第1分钟 (min)	2 摄氏度 (°C)		✓
第2分钟 (min)	2 摄氏度 (°C)		✓
第3分钟 (min)	1 摄氏度 (°C)		✓
第4分钟 (min)	1 摄氏度 (°C)		✓
第5分钟 (min)	1 摄氏度 (°C)		✓
第6分钟 (min)	0 摄氏度 (°C)	水结冰了	确认
第7分钟 (min)	摄氏度 (°C)		确认

温度及现象记录

图3 学生端水温记录模块

(4) 观察解释：学生需要在该模块再次录入温度数据，如图4所示，系统可以根据此数据自动绘制动态温度趋势折线图，帮助学生直观感知水结冰的温度变化全过程。学生基于该可视化图表和实验观察到的现象，提交对“水结冰”现象的科学解释。大模型将基于学生的实验数据及其提交的科学解释进行评价，并给出智能反馈。该模块同样支持多次提交迭代，辅助学生完成从实验现象到科学规律的认知建构。



图 4 学生端观察解释模块

3.2.2. 教师端

教师端设计以“课堂活动导学的调节”为核心目标，旨在通过实时感知与决策支持优化课堂教学流程，示例如图 5 所示。在课堂管控维度，教师端可根据教学节奏动态开启或锁定学生端的活动模块，确保全班探究步调的一致性。在过程监测维度，教师端可以动态监控各组在每一个模块的完成进度、实验记录以及模拟操作等细节，特别是在水温记录模块，教师端可以将所有小组的温度变化数据可视化为一个的温度变化趋势折线图，这不仅为教师提供了精准干预的依据，也为课堂的科学论证提供了直观的数字支持。为了提升教师的决策效率，系统还在各个环节嵌入了基于大语言模型的学情诊断分析与教学建议功能，通过对全班小组提交数据的实时分析，系统能够自动生成小组提交数据的总结与分析，辅助教师迅速捕捉学生的认知偏差。



图 5 教师端面板示意图

4. 研究方法

4.1. 研究对象

本研究选取北京市某小学四年级两个平行班作为研究样本，开展准实验研究。参与学生共计 84 名，授课教师 1 名，按班级划分为实验班与对照班，两班均采取小组协作模式，各划分为 14 个小组，每个组 3 名学生。教学主题选用人教鄂教版科学教材四年级上册《水结冰了》，课程主要以“水结冰了”为主题，引导学生通过小组协作，进行观察、讨论、实验、记录，来探究水在不同温度下的状态变化以及水结冰时的温度。每个小组配备一台平板作为数字化探究的工具。

4.2. 实验流程

本研究采用准实验研究法，在实验班与对照班开展了各为期一课时的教学实践。两班教学均遵循科学探究的五个环节：导入环节、方案设计、装置构建、实验探究、以及总结延伸。

实验班依托本研究设计的智能系统开展教学，学生通过系统的引导和教师讲解共同完成材料选择、模拟实验、水温记录和观察解释这几个活动。实验班总教学用时为 40 分钟，严格执行了标准课时安排，具体教学流程及环节时长如表 1 所示。

表 1 实验班的课堂流程

教学阶段	核心活动内容	环节时长
一、导入环节	复习旧知，引入新课：回顾上节课内容，引出本课探究的水结冰主题。	1 分钟
二、方案设计（活动一）	实验设计与材料选择：教师提出问题如何在课堂探究；学生在平台选择实验材料并填写理由	4 分钟
	交流展示与方案统整：各小组汇报实验材料选择及理由，教师进行统一讲解与纠正。	2 分钟

三、装置构建（活动二）	装置设计与模拟搭建：材料选择完成后，各小组在平台进行“小冰箱”装置的初步搭建与水结冰实验的模拟。	5 分钟
	规范说明与演示：教师提问并讲解模拟实验操作注意事项，并针对数据记录、分工、现象观察等进行实验示范。	6 分钟
四、实验探究（活动三）	动手实验与数据采集：学生领取材料，正式开始实验。每隔一分钟测量并记录水温变化及实验现象。	12 分钟
	数据展示与结果汇报：结合数字化教学平台（折线图）展示各小组实验数据，进行初步交流。	5 分钟
五、总结延伸（活动四）	规律总结与课后拓展：绘制折线图，总结实验结论；教师归纳本课核心知识点，并留下课后思考题。	5 分钟

对照班则采用“通用 AI 工具协同”的模式，其课堂流程与实验班大致相同，唯一区别在于学生在方案设计及总结环节是通过和“豆包”大模型对话来获取建议，装置构建与实验数据记录则通过传统纸质绘制完成。对照班的总教学用时为 47 分钟，超出了预定课时。时间差异主要集中在方案设计与总结延伸环节：实验班凭借平台内置的智能反馈，显著提升了师-生-机交互的效率，促进了教师课堂活动的调节；而对照班的学生在与通用大模型的对话中，需要自行输入背景信息且缺乏针对性引导，导致交互时间延长，教师的课堂活动导学没有得到较好的调节。此外，智能系统的教师端自动生成的折线图与 AI 总结功能有效压缩了数据展示与汇报环节的耗时，体现了智能系统在优化课堂活动导学效率方面的优势。

4.3. 数据收集

本研究构建了全过程、多模态的数据采集体系，旨在通过量化数据与质性数据的相互验证，深度解析智能系统对课堂活动导学与学生探究的影响。在生机交互层面，研究实时提取并汇总了实验班使用系统后的后端日志数据，涵盖了各教学环节学生提交的数据、AI 反馈的内容及学生的修订轨迹等，同时完整留存了对照班与“豆包”大模型的对话记录。通过对这些数据进行初步分析，作为学生科学概念建构质量与实验数据准确性的核心量化指标，用来反映教师课堂活动导学的调节效果。此外，研究对两班全程的授课过程进行了视频录制，并对各小组的协作过程进行了同步录音，以作为后续质性分析的支撑数据。

5. 分析与讨论

5.1. 智能系统对学生探究表现的影响

通过对实验班平台日志数据与对照班豆包对话记录的对比分析发现，基于大模型驱动的智能系统在科学探究的不同阶段均表现出显著的导学优势，两班学生探究表现及交互特征对比如表 2 所示。

表 2 实验班与对照班学生探究表现及交互特征对比

阶段	实验班	对照班
总教学时长	40 分钟 (严格符合课时计划)	47 分钟 (超出预定课时 7 分钟)

活动一：材料选择	平均提交次数 4.07 次，78.6%的小组多次提交假设，其中有 6 个组在收到智能反馈后，将不正确的答案修改为正确的。	平均对话轮数 1.85 轮，材料讨论较为直接，豆包建议后学生通常快速采纳。
活动二：实验模拟	平均操作 86.21 步，存在显著的“试错”与变量控制行为。21%的学生展现出对实验控制更强理解。	
活动三：水温记录	温度范围-9℃~20℃，数据趋势符合物理规律。	
活动四：观察解释	平均提交次数 1.79 次，46.2%的小组在第一次尝试后根据反馈修改了结论解释。	平均对话轮数 1.57 轮，大部分小组仅发送一次消息即结束。

在材料选择阶段，实验班展现了更高质量的决策建构过程。平台日志显示，小组平均提交次数为 4.07 次，78.6%的小组在收到 AI 反馈后多次修订假设，特别是最初选择不安全材料的 6 个小组，在 AI 的针对性引导下均修正为正确选项，并且给出了合理正确的原因解释。相比之下，对照班虽然在豆包的对话过程中成功选择了正确的实验材料，但是该阶段的平均对话轮数仅仅为 1.85 轮，在 AI 给出材料选择的结论之后，只有 3 个组进行了追加提问。大部分组的学生仅仅将豆包视为“标准答案提供者”而非“探究伙伴”。

在实验模拟与水温记录阶段，实验班表现出更强的探究主动性与实验控制力。日志数据显示，学生在模拟实验中的平均步骤数达到 86.21 次，适中的实验重置率反映了学生存在明显的“试错”行为。约 21%的学生展现出深度的实验控制理解，选择先测量基线温度后再进行食盐与冰的混合。在正式实验的水温记录中，实验班的温度范围为-9℃~20℃，数据趋势符合物理规律，证明了系统支持的实验模拟过程具有较好的引导性。而对照班在该活动环节出现了明显的超时现象，学生在使用纸笔绘制及投屏展示的时候，课堂纪律出现了明显的混乱，教师需要花费更多的时间维持。

在观察解释阶段，智能系统的个性化反馈有效促进了认知闭环。实验班中，小组平均提交次数为 1.79 次，46.2%的小组在首次尝试后根据反馈给出的建议修改了结论。而在对照班中，虽然豆包同样可以识别异常数据波动并指出数据和结论不匹配的问题，但由于缺乏结构化的数据输入和反馈环境，几乎大部分小组仅仅发送一次对话就结束了该阶段的讨论。此外，实验班学生的结论解释平均长度约为 60 个字符，且标准差较大，反映出智能系统促进学生思考的差异性，而对照班的学生存在直接照搬豆包答案的情况，这在一定程度上削弱了独立思考的深度。

5.2. 智能系统对教师课堂活动的调节

通过对课堂观察记录及授课视频的定性分析，研究发现智能系统显著增强了教师对课堂复杂环境的感知能力，并优化了课堂活动导学的设计。

在对照班的课堂中，教师需要通过频繁的巡视和随机的学生作品投屏来推测各个小组的进度与认知状态，这种方式浪费了大量的课堂时间。而在实验班中，智能系统教师端的数据看板为教师提供了一体化的监控方式。这种感知能力的增强，让教师能够在课堂上注意到那些在各个活动环节进度滞后的小组，并主动介入指导，减少了学生无意义的时间浪费。此外，

在活动三的水温记录阶段，教师端内置了各组的水温变化趋势图。当个别小组的实验数据出现异常时，系统生成的趋势折线图让教师能够第一时间锁定异常并进行讲解，有助于教师针对性地调整教学策略，为不同小组提供个性化的指导和帮助。最后，大语言模型驱动的智能系统有效压缩了教学时间，促进课堂活动导学的调节。实验班的教学时长要显著优于对照班，这种效率的提升主要源于系统对“数据处理”与“成果展示”环节的自动化支撑。实验班教师不需要花费大量时间指导学生如何绘制实验装置图，也不需要让学生将作品投屏展示，节省了教师维持课堂纪律的时间，同时让学生的注意力始终保持集中。这种由大模型驱动的智能系统支撑的“自动化课堂活动导学”，让教师能够将更多的认知资源投入到高阶问题的启发与科学素养的引导中，从而在标准课时内实现更高质量的教学调节。

6. 结语

本研究针对小学科学课《水结冰了》的设计需求，开发并应用了一套由大语言模型驱动的智能系统。研究表明，该系统通过双端协同架构有效优化了课堂活动导学效能。在学生层面，系统提供的活动导学与智能反馈显著优于通用型大模型，不仅引导学生在规定时间内完成所有学习活动，更通过实时数据反馈帮助小组纠正实验中的认知偏差，实现了深度探究。在教师层面，系统显著提升了课堂活动导学的敏捷性。教师端集成的 AI 总结与数据可视化看板，减轻了教师授课时的认知负担，让其将教学中心转向精准干预与高阶逻辑引导。尽管本研究通过日志数据对对话记录初步验证了智能系统在促进课堂活动导学方面的价值，但仍存在一定的局限性。首先，受限于教学周期与评估维度的侧重，本阶段研究主要聚焦于师生在探究过程中的交互特征与课堂活动导学的效率。在未来的研究工作中，我们将进一步收集并建立两个班级学生的纸质科学活动记录手册，利用内容分析法对学生的推理过程进行更深度的量化评估。

致谢

本研究受到国家自然科学基金重大项目资金资助（项目名称：人工智能赋能教与学的理论与关键技术研究，编号：62293551）、国家自然科学基金委员会信息科学部面上项目（项目名称：基于“批注-反思-研讨”协同阅读教学的过程解构与适应性智能教学策略重构的在线系统研究，编号：62377017）资金资助。

参考文献

- 任长松. (2005). 探究式学习中的科学观教育:18 条内涵. *教育理论与实践*(09), 38-42.
- Dillenbourg, P. (2013). Design for classroom orchestration. *Computers & Education*, 69, 485-492.
- Dillenbourg, P., & Jermann, P. (2010). Technology for classroom orchestration. In *New science of learning: Cognition, computers and collaboration in education* (pp. 525-552): Springer.
- Feng, S., Zhang, L., Wang, S., & Cai, Z. (2023). Effectiveness of the functions of classroom orchestration systems: A systematic review and meta-analysis. *Computers & Education*, 203, 104864.
- Sharma, S., Mittal, P., Kumar, M., & Bhardwaj, V. (2025). The role of large language models in personalized learning: a systematic review of educational impact. *Discover Sustainability*, 6(1), 1-24.

- Van Uum, M. S., Verhoeff, R. P., & Peeters, M. (2017). Inquiry-based science education: Scaffolding pupils' self-directed learning in open inquiry. *International Journal of Science Education*, 39(18), 2461-2481.
- VanLehn, K., Burkhardt, H., Cheema, S., Kang, S., Pead, D., Schoenfeld, A., & Wetzel, J. (2021). Can an orchestration system increase collaborative, productive struggle in teaching-by-eliciting classrooms? *Interactive learning environments*, 29(6), 987-1005.