

多模态数据驱动的 STEM 项目式学习过程评估系统构建与应用

Construction and Application of a Multimodal Data-Driven STEM Project-Based Learning Process Assessment System

谢颖¹, 乔爱玲^{1*}

¹ 首都师范大学教育学院

* qiaoal@126.com

【摘要】 针对 STEM-PBL 评价实践困境, 本研究计划开发并验证一套学习过程评估系统。采用基于设计的研究范式, 沿“理论构建、原型开发、实证检验”路径展开。首先, 构建涵盖认知、行为、社会与元认知维度的评估指标框架。继而开发集成多模态数据采集、融合分析与可视化反馈的原型系统, 运用认知网络分析等方法, 实现对学习协作与设计过程的自动化解析。最后通过准实验与案例研究, 在真实课堂中检验系统有效性, 并探究其对学习投入、协作质量与元认知的影响机制。本研究计划为 STEM-PBL 过程性评价与反馈提供整合性方案, 也为技术赋能的动态评估提供实践参照。

【关键词】 STEM-PBL; 学习评估; 多模态数据; 学习分析

Abstract: This study plans to address practical challenges in STEM-PBL assessment by developing and validating a learning process assessment system through design-based research. Following a "theoretical construction—prototype development—empirical testing" sequence, it first establishes a multi-dimensional assessment framework around cognitive, behavioral, social, and metacognitive dimensions. A prototype system is then built to integrate multimodal data collection, fusion analysis, and visual feedback, applying methods like epistemic network analysis to automatically parse collaborative and design processes. Finally, classroom quasi-experiments and case studies evaluate the system's effectiveness in enhancing learning engagement, collaboration, and metacognition. This research plan to offer an integrated solution for formative assessment and feedback in STEM-PBL, along with actionable insights for technology-supported dynamic evaluation.

Keywords: STEM-PBL, Learning assessment, Multimodal Data, Learning analytics

1. 研究目的

针对 STEM 项目式学习 (STEM-PBL) 中长期存在的评价理论与实践脱节问题, 即传统评价方式难以有效衡量复杂问题解决过程中的高阶思维、协作与实践能力 (Gao et al., 2020), 本研究计划构建、实施并验证一个内嵌于 STEM-PBL 学习过程的多模态动态评估系统。在理论层面, 整合学习科学、表现性评价与多模态学习分析的相关成果, 构建能够刻画认知、社交、行为与元认知等多维能力发展的过程评估模型, 以弥合宏观教育目标与微观学习证据之间的鸿沟。在技术与方法层面, 研发多模态数据融合分析引擎与可视化反馈系统, 通过对非结构化过程数据的自动化解析, 将“协作质量”等抽象概念转化为可量化的计算指标。在实践层面, 于真实课堂环境中检验该系统的教学有效性及用户体验, 探究其如何通过实时、可视化的过程性反馈影响师生的元认知与教学决策, 最终形成可直接支撑 STEM-PBL 教学的评估工具与设计原则。

2. 研究背景

作为开展科学、技术、工程、数学多学科于一体的跨学科综合教育，STEM 教育倡导通过基于科学和数学的概念和程序解决问题，结合工程策略和应用适当技术，对于发展学生的高阶思维能力、人际交往能力以及创造能力等核心能力具有重要作用（Sánchez Carracedo et al., 2018）。其中 PBL 与 STEM 教育的结合（STEM-PBL）可以提供“情境化的真实经验，使学生能够构建有意义的、强大的科学、技术、工程和数学概念”（Capraro & Slough, 2013），成为全球备受欢迎的教学模式。然而，其评价体系的发展滞后于 STEM 教学实践的创新，陷入三重困境。

其一，评价内容与学习目标错位。标准化测试或作品评分难以评估项目进程中动态发展的核心能力（Gao et al., 2020），如迭代设计中的决策推理、跨学科知识整合及从失败中学习的学习韧性。其二，评价方法在信度、效度与可操作性上存在局限。教师观察或学生自评主观性强且易产生回顾性偏误（Hsiao et al., 2022）；而学习管理系统日志分析多聚焦于点击流、学习时长等表层行为（Xia & Qi, 2024），难以触及认知与社会性互动等深层内涵。其三，技术赋能评价的潜力未被充分挖掘。当前技术应用多为零散的、事后分析式的工具（Karakaya & Yilmaz, 2022），缺乏对多源异构数据同步采集、融合分析与即时反馈的完整架构。

基于上述分析，本研究以多模态学习分析为核心理论视角与方法论框架。学习本质上是具身化与情境化的，其过程证据广泛蕴含于语言、行为、生理等多模态数据流中。通过系统性采集并计算融合这些数据，能够形成对复杂学习过程更为全面、客观且精细的动态表征。因此，本研究致力于解决的核心问题是**如何设计并实现一个基于多模态学习分析的技术系统，以自动化、可视化的方式对学生的 STEM-PBL 能力表现进行过程评估，从而有效提升教学与学习的质量与效率？**本研究拟定提出一种创新的评价方案，为技术赋能的动态学习评估提供实践参照。

3. 研究方法

本研究拟采取“理论构建-原型开发-实证检验”的路径。第一阶段为理论框架构建（图 1），通过系统的文献综述确立理论基础，随后对一线 STEM 教师及学科专家进行深度访谈与焦点小组讨论，获取实践需求。在此基础上，采用德尔菲法组织专家咨询，对初步构建的评估维度与指标进行评议，确立具有共识的多模态 STEM-PBL 过程性评估框架。

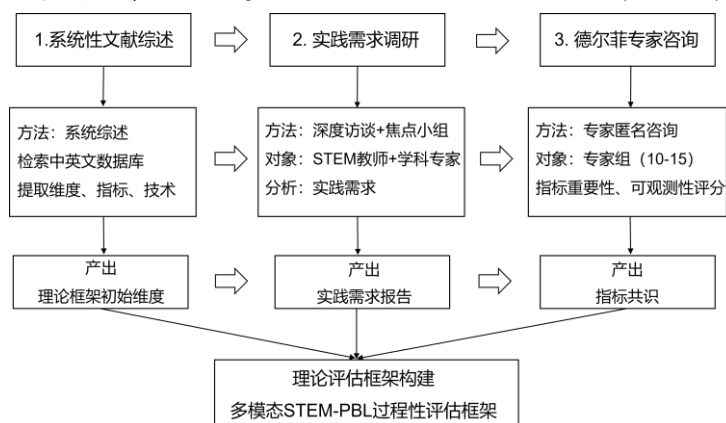


图 1 理论框架构建流程图

第二阶段为系统原型开发（图 2），核心任务是将理论框架转化为可运行的技术原型。为此，本研究设计了包含“数据采集、处理分析、可视化反馈”三层结构的系统架构。在数据采集层，系统整合定向麦克风、全景摄像机与学习管理平台的 API 接口，在 PBL 任务执行过程

中同步捕获音频、视频及数字制品日志，形成多模态数据流。处理分析层则对上述数据进行综合处理。运用自动语音识别与自然语言处理技术，对小组对话进行转写与语义编码，再通过认知网络分析刻画协作对话中的概念关系结构，从而量化评估 STEM-PBL 学习过程的关键维度；其次，为捕捉协作中的非言语互动，系统引入基于 MediaPipe 库的计算机视觉方法，通过实时姿态估计提取成员的身体朝向、手势与间距等特征，进而计算“小组空间凝聚度”与“任务协同焦点时间比”，以客观反映协作参与模式与注意力分布。此外，采用滞后序列分析识别设计迭代中的时序依赖模式，对应元认知调控行为。三类数据经毫秒级时间对齐后融合为统一事件日志，通过跨模态事件关联实现因果推断。数据处理采用近实时批处理，教师仪表盘每 10 分钟刷新一次，以平衡反馈及时性与课堂流畅性。在可视化反馈层则基于 Web 技术开发交互式教师仪表盘，将上述分析结果以时间线、网络图、热力图等形式动态呈现，为教学干预提供直观依据。

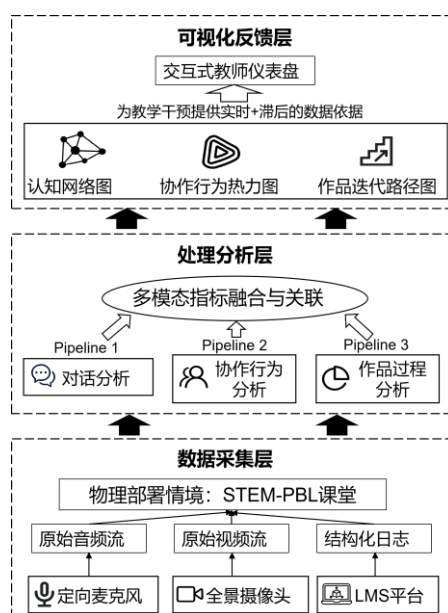


图 2 系统原型开发流程图

第三阶段为实证检验，在中学真实情境中开展两轮迭代研究。参考基于设计的研究范式，首轮迭代旨在对系统原型进行形成性评估与优化，核心是确保工具的可用性与可靠性。首轮迭代在一个初中 STEM 项目班中展开，采用单组案例研究设计，全面部署并测试系统原型（V1.0）。数据收集聚焦于技术可行性与初始用户体验，基于初步测试结果，重点识别出原型系统在算法稳定性、用户交互逻辑及教学场景适配性等方面存在的主要不足，改进并形成 V2.0 系统版本，为后续开展严格的效能实证研究奠定了可靠基础。次轮迭代旨在对优化后的系统进行效果验证与影响机制探索，核心是检验其教育有效性并深化理论理解。采用准实验设计，同步采集多模态过程数据，两组教师保持相同的干预频率与时长，唯一差异在于实验组教师基于仪表盘数据进行诊断，对照组教师基于常规观察进行诊断。收集前后测量表、过程性多模态数据、作品成绩及访谈资料。基于此轮分析，研究最终形成完善的评价模型、定型的工具系统，并总结出可迁移至同类情境的 STEM 学习过程评价设计原则与实践指南。

4. 研究阶段与初步成果

目前，本研究处于正式实施的启动与奠基阶段。前期构想与规划已完成，并已转化为一个具体、可操作的研究方案，为核心工作的开展做好了全面准备。在研究整体规划与设计方面，已经完成了核心研究问题的凝练并确立了以“基于设计的研究”为核心，融合混合方法的总体研究范式。研究的技术路线已经明确，即通过构建多模态数据采集与分析系统，实现对

STEM-PBL 学习过程的动态评估。包含“理论构建、原型开发、实证检验”的研究路线图已绘制完成，并完成了对初步合作研究场域的联络与沟通。

其次，在理论基础与技术准备方面，已经完成了初步的探索性工作。研究已围绕“多模态学习分析”“STEM 过程性评价”等主题进行了广泛的文献调研，确定了以学习科学、学习过程评价和社会文化理论作为本研究的核心理论视角。同时，已对关键技术的可行性进行了初步探索与测试，例如利用现有开源工具（如 MediaPipe）尝试了课堂视频中人体姿态的初步提取，并使用在线语音识别服务测试了对话转写的效果等。

5. 未来研究工作

接下来将有序推进第一阶段至第二阶段的工作，具体计划如下。第一步，完成理论评估框架的构建与验证。通过系统文献综述及对 15-20 名 STEM 教师与学科专家的访谈调查，明确评价需求。整合文献与实证数据，形成多模态评估指标框架。随后邀请教育学、学习分析、STEM 教学等领域专家（约 10-15 人）开展两至三轮德尔菲咨询，通过统计分析各指标的重要性与可观测性评分，确立具有良好内容效度的正式评估框架。预计此步骤持续 3-4 个月。第二步，启动多模态分析系统的原型设计与开发。在框架基本确定后，进行系统架构设计，明确数据流与技术选型，并依据指标进行算法模型选择与适配。同步开展前端仪表盘原型设计，与教师进行可用性测试迭代，目标是完成核心功能可演示的集成原型系统 V1.0，为第一轮实证循环做好准备。第三步，开展第一轮实证循环。在合作学校选取一典型 STEM 项目，在班级中部署系统 V1.0。本轮关注形成性评估，收集系统技术性能数据、师生用户体验反馈及多模态数据初步分析结果，检验指标计算的有效性。通过分析识别技术瓶颈、交互设计问题及教学融合障碍，为开发更稳定易用的 V2.0 系统提供改进依据，并同步完善下一轮效果验证的研究设计。

参考文献

- Capraro, R. M., & Slough, S. W. (2013). Why PBL? Why STEM? Why now? An introduction to project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach. In R. M. Capraro & S. W. Slough (Eds.), *STEM Project-based learning: An integrated science technology engineering and mathematics (STEM) approach* (pp. 1-6). Rotterdam, The Netherlands: Sense.
- Gao, X., Li, P., Shen, J., & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education*, 7(1), 24.
- Hsiao, J. C., Chen, S. K., Chen, W., & Lin, S. S. (2022). Developing a plugged-in class observation protocol in high-school blended STEM classes: Student engagement, teacher behaviors and student-teacher interaction patterns. *Computers & Education*, 178, 104403.
- Karakaya, F., & Yılmaz, M. (2022). Teachers' views on assessment and evaluation methods in STEM education: A science course example. *Journal of Pedagogical Research*, 6(2), 61-71.
- Sánchez Carracedo, F., Soler, A., Martín, C., López, D., Ageno, A., Cabré, J., ... & Gibert, K. (2018). Competency maps: an effective model to integrate professional competencies across a STEM curriculum. *Journal of Science Education and Technology*, 27, 448-468.
- Xia, X., & Qi, W. (2024). Driving STEM learning effectiveness: dropout prediction and intervention in MOOCs based on one novel behavioral data analysis approach. *Humanities and Social Sciences Communications*, 11(1), 1-19.