

多模态智能体支持下的跨学科情境学习：理论框架与机制建构

Multimodal Agent-supported Interdisciplinary Contextual Learning: Theoretical Framework and Mechanism Construction

王晗¹, 吴永和^{2*}

上海市嘉定区青少年科创集散地

华东师范大学

* yhwu@deit.ecnu.edu.cn

【摘要】 针对传统情境化学习“情境静态化”导致问题解决流于形式的困境，本研究借助多模态智能体的全景感知能力，提出“情境触发学习”概念，强调情境应作为主动驱动学习的交互主体。基于认知负荷与最近发展区理论，研究构建了多模态智能体支持下的跨学科情境学习理论模型。该框架通过动态研判介入时机，严格遵循必要时介入、最小化支架的原则，同时引入元认知与社会性支持，实现学习支持与学习评价的闭环调节。本研究明确了智能体的功能边界，为智能技术深度赋能跨学科情境化学习提供了系统性的理论模型与机制参考。**【关键词】** 多模态智能体；情境触发学习；跨学科情境；学习支架；动态适应

Abstract: In view of the dilemma that the traditional contextual learning "static context" leads to problem solving becoming a mere formality, this study proposes the concept of "Context-Triggered Learning" with the help of the panoramic perception ability of multimodal agents, emphasizing that the context should be the interactive subject of active driven learning. Based on the Cognitive Load Theory and the Zone of Proximal Development (ZPD), a theoretical model of interdisciplinary contextual learning supported by multimodal agents is constructed. By dynamically studying and judging the timing of intervention, the framework strictly follows the principles of intervention when necessary and minimizing support, and introduces metacognition and social support to realize the closed-loop adjustment of learning support and learning evaluation. This study defines the functional boundaries of agents, providing a systematic theoretical model and mechanistic reference for deeply empowering interdisciplinary contextual learning through intelligent technology.

Keyword: Multimodal Agents; Context-Triggered Learning; Interdisciplinary Context; Learning Scaffolding; Dynamic Adaptation

1. 前言

在当今教育背景下，情境化学习已超越单纯的教学策略，成为促进学生思维能力和创新能力发展的重要教学范式。教育部等七部门发布的《关于加强中小学科技教育的意见》明确提出，应基于真实情境问题解决培养学生的跨学科素养与科学思维(教育部等七部门, 2025)。这深刻揭示了情境化学习与跨学科学习之间的天然同构性。情境化教学强调创设学习情境的氛围，引导学习者自觉主动学习(李吉林, 2016)，是推动课堂教学改革的重要支点。然而，

在跨学科教学中，情境往往被用作课程背景，未能持续介入学习过程。教师在现实课堂中通常难以对多名学习者的认知变化进行同步感知，也难以及时在学习受阻的关键节点提供个性化支持。人工智能技术在教育领域的不断深入，一定程度上缓解了资源获取与反馈效率的问题，但其支持方式多以被动响应为主，依赖学习者主动提问或按照预设路径推进学习活动。由此，如何将作为背景的情境转化为能够持续驱动学习发生与演进的作用机制，成为情境化学习研究亟需解决的问题。

近年来，随着人工智能技术在感知、理解和交互层面的发展，催生了具备多模态感知与动态响应能力的智能体。这类智能体能够持续捕捉学习者在任务推进过程中的行为轨迹与认知状态，这使情境从静态呈现的学习背景转化为参与学习调控的动态要素成为可能。

基于上述背景，探讨多模态智能体如何支持跨学科情境中学习的触发、调节与推进，对于深化情境化学习理论并回应人工智能时代的教学设计需求，具有显著的研究意义与实践价值。本研究尝试在跨学科教学情境中引入多模态智能体，围绕“情境如何触发学习”这一关键问题，构建情境触发学习的理论设计框架，并系统分析智能体在学习支持机制中的作用定位与运行逻辑。

2.情境学习的理论转向与研究问题界定

2.1. 情境化学习的实践困境与理论转向

2.1.1 从情境化学习到情境触发学习

情境化学习理论主张学习者围绕特定主题与情境展开充分互动，在互动中实现经验的生成与拓展（于泽元、那明明，2023）。理想形态的情境化学习应依循学习者的认识过程，适时提供有目的、有针对性的情境化支持策略（陈春芳，2022）。然而，实际教学实践中，情境在教学系统中的作用定位与其理论设想之间存在一定落差。这种落差并非单纯的教学设计失当，而是反映出情境化学习在机制层面的内在困境。就学习过程而言，跨学科情境中的学习过程具有不确定性，学习者须在理解偏差、策略选择与认知冲突中不断调整行动路径。就教学组织而言，课程设计高度依赖预设情境和既定流程，对学习过程中的动态变化，难以满足学习者的个性化发展诉求。因此，当学习者在真实问题解决中出现学习停滞或策略偏移时，静态的情境本身无法提供新的反馈，教师也难以在动态、变化的教学情境中提供针对性、即时性的学习支持。

人工智能技术的发展对高阶思维与复杂问题解决能力的培养需求进一步放大了情境与认知之间的矛盾。学习者需要在学习过程中围绕教学主题产生更有深度的思考，而恰当的问题情境能够有效激发学生的深刻思考（朱荣光，2021）。这一现实要求情境化学习完成从背景化向触发式的理论转变：情境不再只作为有待理解的外部条件，而应成为一种能够随学习进程变化、持续介入认知发展，在学习关键节点对学习者的理解与行动产生实质性影响的交互介质。

基于上述理论转向，本研究提出情境触发学习作为分析情境在学习系统中运行方式的核心概念。本研究对情境触发学习做出形式化定义：情境触发学习是具备动态交互性、节点触发式、全程伴随性三类属性的学习过程，即在具有真实性、复杂性与指向性的情境中，通过对学习过程中关键认知事件的感知与解释，引发学习者的认知失衡与意义建构需求，驱动其主动开展学习活动，并在持续的情境与学习者双向互动中实现知识生成、经验重组与能力发展。不同于传统意义上的问题背景与问题情境，此类概念多作为预设的静态内容，仅在学习初始阶段承担任务导入功能，属于学习活动的外部条件；而情境触发学习将情境视为贯穿学

习全过程的动态交互主体，强调其对学习进程的持续性伴随与调节作用。不同于基于问题的学习，该模式以固定的真实问题驱动自主探究，情境在探究过程中不发生变化，学习者受阻时依赖自主摸索或教师被动介入；而情境触发学习聚焦学习过程的微观动态，强调情境可依据学习者的实时状态主动调整，在特定认知事件触发时产生适配性干预，实现对学习进程的驱动。结合本研究的技术载体与应用场景，进一步将多模态智能体支持下的情境触发学习界定为：多模态智能体依据学习者所处跨学科情境中的实时行为表现与认知状态变化，在学习关键节点通过改变情境参数、提供适应性支架等方式，重构学习者的认知平衡，推动深度学习持续发生的学习过程。

2.1.2. 情境触发学习的实施流程

为构建智能体支持的情境触发学习实施流程，本研究对情境感知、精准决策及人机协同等领域的经典文献进行梳理，以厘清各环节的理论基础与技术路径。

情境感知是智能系统提供支持的起点。黄国珍等（2008）在情境感知泛在学习研究中指出，系统运行的基础是感知识别学习者在真实世界中的状态与环境参数。杨现民与余胜泉（2015）将智慧教育架构划分为感知层与计算层，强调多模态技术在行为数据捕获中的功能价值。如何将感知到的情境数据转化为决策，牟智佳与武法提（2015）指出系统应基于多源数据判断学习者的个人特征，并基于学习者模型提供个性化学习支持服务。卢宇等（2023）强调生成式 AI 具备对话式情境理解和协作式动态支架支持与反馈能力，能够提供个性化的启发式支架与社会性交互。为了维持系统的持续适应性，Clow(2012)提出“五步学习分析模型”，强调将评价结果实时回流至系统以触发干预，从而优化后续决策。

综上所述，已有研究虽然分别验证了感知、决策或评价等独立环节的有效性，但尚未形成“最近发展区（ZPD）判断”与“多模态生成式支架”两者结合的完整闭环模型。基于此，本研究汲取前人成果，将情境触发学习的运行流程重构为六个环节（见图 1）：

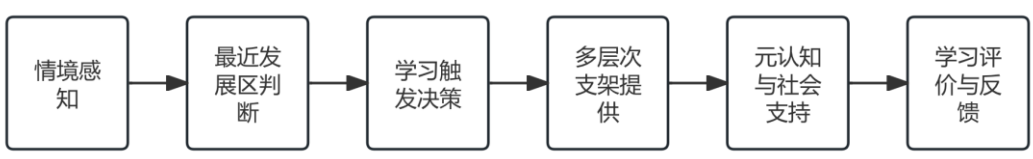


图 1 情境触发学习的六大环节

该流程在以下维度实现了机制创新：第一，在数据整合层面，摒弃单一的数据源依赖，转而通过动态计算整合多源数据，建立高还原度的学习情境映射；第二，在认知研究和判断层面，系统通过认知状态模型判断学习者的最近发展区（ZPD），以确保教学内容的难度始终维持在最佳区间；第三，在触发机制层面，决策机制并非时刻在线，而是通过识别学习停滞、错误尝试或认知犹豫等关键事件来激活。第四，在支架供给层面，遵循必要时介入、最小化供给，生成自适应的跨学科支架。第五，在高阶支持层面，系统还融合了指向高阶思维的元认知与社会性资源支持，并借助多维评价数据形成闭环，持续推动智能体模型的迭代优化。

2.2. 现有智能技术支持体系的局限与理念转向

近年来，智能技术赋能学习已成为教育研究的重要趋势，主要体现为两个紧密关联的方向：多模态学习分析与自适应学习系统。多模态学习分析侧重于通过整合文本、语音、视觉、生理及行为等多源异构数据，实现对学习过程的全面感知与深度诊断，其核心在于数据驱动的全景描述与评价优化（李嘉瑶，2024；张艳明等，2023）。与此同时，自适应学习系统则侧

重于路径适配，即通过知识状态诊断与学习者模型构建，实现学习内容与路径的个性化推送（许琼，2024；郑金明，2025）。其核心理念是基于历史表现数据的、后验的路径调整，以匹配学习者当前的认知水平。

现有技术体系在支持跨学科情境化学习时仍然存在显著局限：一是情境静态化与介入被动性。现有多模态分析虽能全面描述学习过程，但多数将情境视为固定的数据分析背景，分析结果往往用于事后评价与反馈，未实现对学习过程的动态介入；自适应系统则依赖预设式知识图谱与问题，通过人机交互的用户界面为学习者提供学习支持（王萍等，2018），其核心逻辑路径适配，本质上是在诊断出学习者知识缺陷后，直接推送预设的资源以降低学习难度或缩短学习路径，以预设的最优路径替代学习者的自主探究，既缺乏对跨学科任务中不确定性问题的及时响应，也与本研究情境触发所秉持的、通过情境动态调整驱动学习者自主建构的理念存在本质差异。二是支持机制缺乏教育性节制。现有系统往往倾向于提供泛在的支持，多模态分析追求全时监控，自适应系统致力于路径推送，加之当前生成式人工智能提供的支架过于直接或密集，极易引发认知外包——学习者将高阶思维让渡给算法，使分析能力弱化（Grassini, 2023）。对完整解决方案的过度依赖，不仅诱发学习者在面对不确定性问题时出现习得性无助，更可能导致其独立构建认知路径能力的削弱（Dwivedi et al, 2023）。三是跨学科认知融合不足。现有系统多针对单一学科或明确技能的结构化训练，其数据模型与推送逻辑难以捕捉和支撑跨学科情境中复杂问题解决的认知跳跃、策略转换与概念整合过程。

从认知建构主义的视角来看，深度学习的发生源于现有认知与新经验之间的不平衡。因此，情境触发学习的核心并非优化或者缩短学习路径，而是在感知到学习者处于不顺利状态时，通过动态改变情境参数来主动重构当前的问题。与现有系统的根本区别在于，本研究构建的多模态智能体支持的跨学科情境学习框架，实现的理念转向可概括为三个维度。从预设路径适配转向动态情境触发，突破预设路径的限制，在认知停滞或策略固化等节点，改变情境参数，驱动深度学习。从全时泛在支持转向最小化节制介入，秉持必要时介入与最小化支架的原则，将支持严格限定于最近发展区内的关键节点，防范认知外包。从终结性结果评价转向过程性闭环迭代，构建评价闭环，使评价数据实时回流，驱动智能体触发策略与情境设计的持续迭代。

概括而言，多模态智能体具备整合多源信息的能力，展现出较高的通用性与泛化性（卢宇等，2023）。本研究并非对多模态数据或自适应技术的简单整合，而是针对跨学科情境学习的动态性、生成性和教育性本质，实现技术可能性向教育实效性的转化，构建一套兼顾技术精度与育人温度的介入框架。

2.3. 研究问题界定

基于上述理论分析与文献梳理，本研究致力于构建多模态智能体驱动的跨学科情境学习框架。围绕情境触发学习的运营机制，本研究聚焦以下三个层面的问题：

第一，感知的精度与时机。智能体如何在跨学科情境中重构出高还原度的学习者认知状态，为介入行为确立触发点。

第二，介入边界与尺度。智能体如何设计具有节制性的动态支架，在提供技术脚手架的同时，守护学习者的探究主权。

第三，反馈的深度与闭环。智能体如何建立基于过程行为的深层分析，使评价数据不仅作为结果，更作为反向校准情境参数与持续迭代的动力源。

上述三个问题分别指向情境触发学习中感知、介入与反馈三个关键环节，依次回应智能体：何时介入、何时支持、怎样改进的运行逻辑，构成本研究的基本框架。

3.多模态智能体支持的跨学科情境触发学习设计框架

3.1. 多模态智能体载体形态、多模态界定与师生机协同定位

在情境触发学习框架中，多模态智能体在形态上并非孤立的问答式 Chatbot，而是内嵌于中小学跨学科项目式学习专用网络平台的伴随式智能助教模块，与平台内的任务发布、过程记录、成果提交、师生互动功能完全打通，无需学习者切换独立应用即可实现全流程伴随支持。在多模态的数据层级上，它具体融合三个维度。操作行为模态，对应平台交互记录、鼠标轨迹与停留时间等数据，用于识别学习者的操作停滞、错误重复、策略切换等行为特征。视觉模态，对应通过设备普通摄像头捕获的视线焦点与面部微表情数据，用于识别学习者的专注度、困惑、焦躁等情绪与认知状态。语音交互模态，对应任务场景内的师生对话、小组讨论音频转写内容，用于识别学习者的认知卡点、概念误解与协作困境。目前，国内外主流的大语言模型均可实现以上多模态数据的获取和分析，本研究的机制设计均基于现有模型的成熟能力展开。

基于上述载体与数据感知，多模态智能体在情境触发学习中主要承担以下三重角色：第一，情境变化的感知者。智能体辅助获取学习者与任务、环境之间的交互变化。第二，介入决策者，智能体根据学习进程中的情境变化，决定是否需要触发学习支持。第三，支架供给者。智能体作为触发机制的执行单元，在不破坏学习主体性的前提下，对学习进程施加适度干预，打破既有认知平衡，促使学习者重新审视问题、调整策略。

在该框架中，智能体的所有运行动作均受控于教师，教师的主体作用贯穿学习全流程。教师负责宏观跨学科情境的创设与核心问题设计，为学习活动划定整体边界。教师可自主设定智能体介入的规则阈值，包括允许学生试错的最长时间、单任务最大介入次数、支架提示的信息密度上限等，适配不同学段、不同任务的教学需求。教师可随时查看智能体的感知数据、触发记录与支架内容，对智能体的判定偏差进行人工校准。当智能体多轮干预无效，或学生遭遇涉及深层情感与价值观的困境时，系统自动将干预权移交回教师，由教师提供社会性支持与情感疏导。由此，构建学生主导探究、智能体微观伴随、教师全流程把控的三元协作体系。

3.2. 智能体的多模态情境感知与学习者发展区判断

在情境触发学习框架中，智能体的感知能力不再局限于教学进度的线性追踪，而是转向对学习者的全面认知状态的计算性推断。这一过程遵循“数据感知-状态映射-区间判定”的逻辑闭环（见图2）。

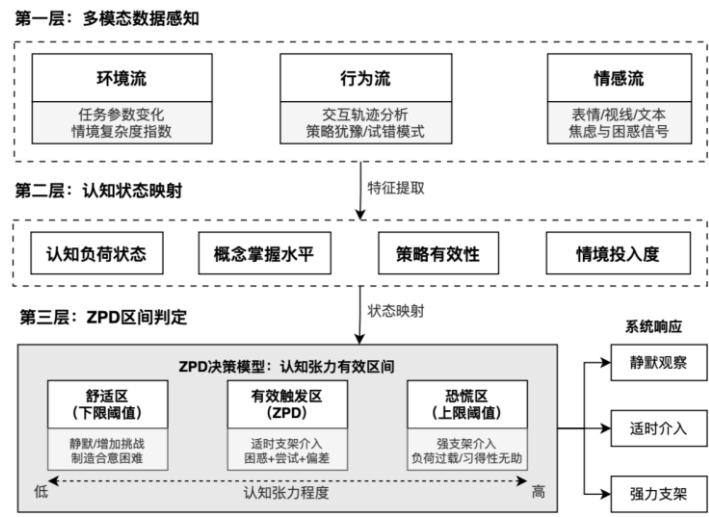


图 2 智能体多模态情境感知与 ZPD 判定机制框架

本研究的感知与判定机制建立在现有情感计算与多模态学习分析的实证数据基础之上：第一，在数据感知与状态映射环节。普遍关注的是机器能否真的看懂学生的认知状态？现有的实证研究已给出肯定答复。张家华等（2022）的研究表明，通过融合外显行为、心理及生理等多模态数据，能够以极高准确率表征学习者实时的认知负荷与情绪状态。此外，Blikstein 与 Worsley（2016）的实证亦证实，利用面部动作编码系统与姿态追踪等技术进行多模态整合分析，可有效且无感地识别学生是处于积极沉浸还是困惑迷茫。本框架正是吸收了这些已被验证的感知参数，将抽象情境转化为可计算的指标。第二，在最近发展区预判环节。本框架将抽象的最近发展区具象化为可计算的触发阈值区间：舒适区（轻松作答）、有效发展区（伴随适度认知摩擦与策略改进），以及恐慌区（重复错误与受挫焦虑）。此区间的量化同样有赖于实证支撑。D'Mello 等（2014）关于情绪动力学的经典实验表明，通过追踪复杂问题解决中的错误序列、停滞时长与困惑表情，可精准区分学生处于有益的认知挣扎还是无效的受挫。本研究的最近发展区判定即基于此实证逻辑：智能体在学生处于舒适区时静默观察；仅当多模态数据研判其认知过载、陷入无效受挫时，才判定触及最近发展区上限，进而精准激活干预机制。

综上所述，上述判定机制以多模态感知为基础，确保其始终遵循必要时介入、最小化供给，为后续情境触发与学习支持提供合理的介入时机，而非取代学习者的判断或教师的专业决策。

3.3. 跨学科情境触发学习的理论设计框架

基于前述智能体的功能定位与感知机制，本研究构建跨学科情境学习的理论框架（见图 3）。该框架将“感知-判断-决策-支持-反馈”整合为一个能够实现对学习者认知路径的伴随式引导与动态调节。

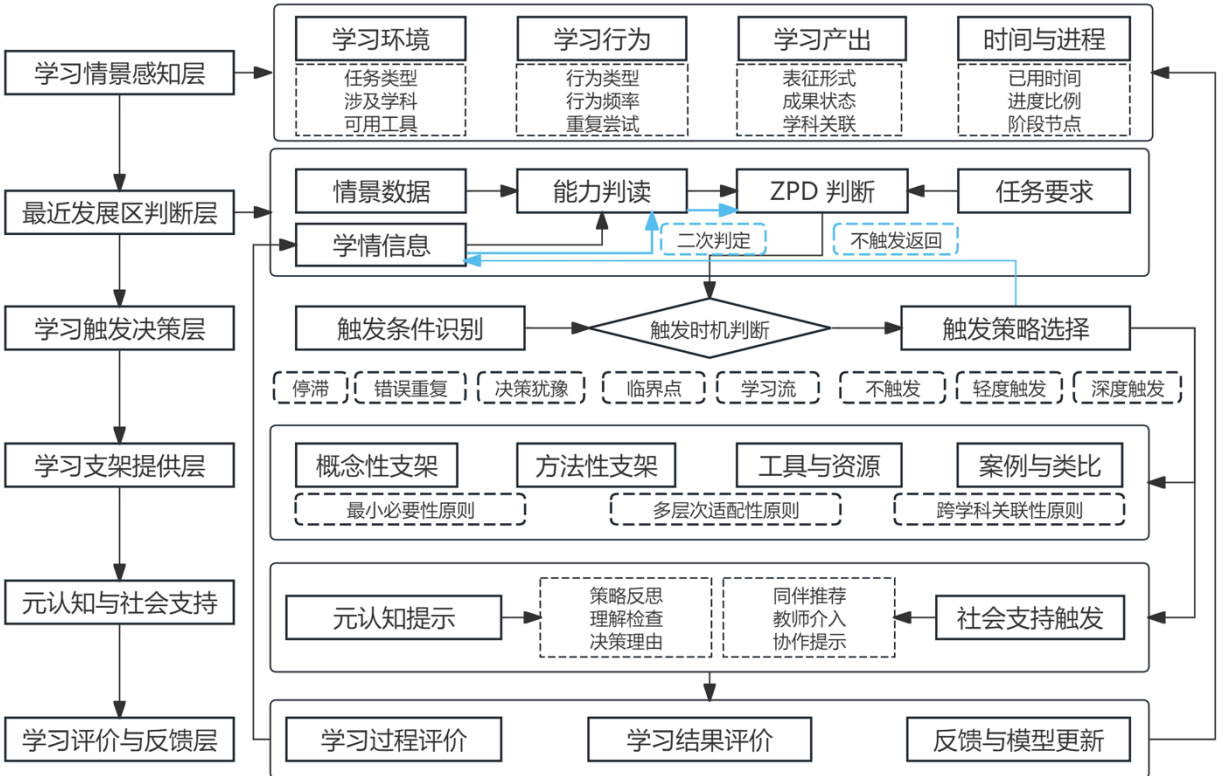


图 3 多模态智能体支持下的跨学科情境学习理论模型

从内在运作上看，该框架主要呈现了智能体在教学过程中的递进式支持路径：首先，情境感知与最近发展区判断是为了读懂学生。系统通过学习环境、学习行为和学习产出等数据的采集来推断学生当前真实的认知难点。其次，学习触发决策基于特定认知事件（如学习停滞或错误循环）进行时机研判，决定是否激活干预机制，若无法触发则说明最近发展区判断出现偏差，需要返回学情信息对最近发展区进行二次判定。最后，学习支架提供层和元认知与社会支持层共同构成了分层引导体系，根据认知障碍的具体类型不仅匹配概念性、方法性或工具性支架，更在学习者陷入深度困境时引入反思提示或同伴/师生协作。而底部的学习评价与反馈层，将过程数据提炼为素养画像，并将数据回流至感知层，用于不断优化智能体下一次的介入。

3.4. 理论框架的研究设计与验证路径

然而，上述框架仍停留于理论建构层面，其在真实学习情境中的适用性与作用机制仍有待进一步检验。基于此，本研究采用基于设计的研究方法作为核心研究方法，制定三轮迭代的实证验证计划，具体路径如下。

在变量的操作化映射与预实验设计方面，框架的运作被解构为干预、过程、结果的观测路径。以初中跨学科 STEM 项目学习为应用场景，设置实验组与对照组，实验组采用本研究构建的多模态智能体支持框架，对照组采用传统静态情境教学模式，其中智能体的支持方式构成干预条件；学习者的最终问题解决质量与跨学科整合能力作为结果表征；认知停滞事件频次、支架触发与响应间隔及试错频率等作为过程性观测指标，以此验证框架在动态干预过程中的敏锐度。在多模态数据的机理挖掘与三轮迭代方面，验证该框架需突破单一的成绩测验。每一轮迭代均深度融合系统交互记录与视线、表情等多模态数据，在时序与结构双重维度上，审视学习者在遭遇智能体重构情境或延迟支架供给后，是否发生策略调整与深层意义建构的转移，从而在数据层面印证感知、触发、反馈闭环的传导有效性；每一轮迭代结束后，

基于数据结果优化智能体的触发阈值与支架设计，开展下一轮验证。在边界条件的探寻上，验证焦点指向最小化介入原则，通过前后测对比与半结构化访谈，考察智能体设置的合意困难是否能在促进理解深化的同时，防范认知外包与习得性无助。

明确上述验证逻辑与实证路径，为本框架的科学性落地提供了方法论依据。而要使这些验证设想具备可运行的技术实体，需进一步厘清智能体在各个关键节点的介入规则。据此，本研究将详细展开该框架的核心支持机制设计。

4.情境触发学习中的智能体支持机制设计

4.1. 学习触发机制：介入时机的识别与介入原则

在本研究建构的理论框架中，“触发”是一种基于计算性状态判断的策略性干预，智能体须在认知外包与习得性无助之间审慎维持平衡。为此，本研究确立以特定认知事件识别为触发逻辑，以必要时介入为算法约束的基本机制。

4.1.1. 触发时机：三类特定认知事件的识别

区别于传统教学仅关注结果正误，情境触发学习强调对学习过程微观状态的捕捉。本研究界定了三类具有高触发价值的特定认知事件：一是学习停滞，多模态数据（如长时间的操作静止、无效的工具点击）显示学习者在当前任务节点无法推进；二是错误尝试循环，学习者重复提交同一类型的错误方案，或在尝试不同策略后迅速回退到原错误路径；三是决策迟疑，学习者在特定决策节点表现出高频的鼠标悬停、策略切换或面部困惑表情，处于新旧知识的冲突地带。

由于单一模态数据易产生解释歧义，本系统依托多模态学习分析技术进行交叉验证。为增强机制的可观测性与后续实证研究的可操作性，本研究构建了多模态数据与底层认知状态之间的交叉映射。为直观呈现该映射逻辑，表 1 以跨学科 STEM 任务“设计环保水净化工程装置”为例，具象化地阐释了三类认知事件的研判依据。需要特别说明的是，不同学科与任务情境下的具体研判标准可灵活调整，但多模态联合表征以推断认知状态的底层运算逻辑具有跨情境的普适性。

表 1 多模态数据与特定认知事件的特征映射
(以“环保水净化工程装置设计”任务为例)

认知事件	认知状态推断	通用多模态映射特征	具体任务案例表征
学习停滞	认知负荷过低 (走神)或极度 超载(迷茫无措)	行为：长时间静止或无意义点击。	停留在过滤材料选择界面超过预设时长(如 3 分钟)未拖拽任何材料入槽；视线未聚焦于任何核心组件；身体呈被动后仰姿态，陷入操作停滞。
		视觉：目光涣散，视线游离于屏幕外。	
错误循环	陷入认知盲区； 策略失效	姿态：身体后倾，伴随长时间沉默。	连续 3 次将细沙置于最上层导致水流堵塞，每次报错后迅速撤销并固执重复该错误排列；视线直盯报错区；伴随焦躁的鼠标重击。
		行为：极短时间内高频重复同一错误动作。	
		视觉：直盯特定区域，眉头紧锁。	
		姿态：鼠标点击力度加大，表现急躁。	

决策迟疑	处于新旧知识冲突地带；面临多策略选择犹豫	行为：鼠标在多项间高频悬停，无实质确认。	鼠标在活性炭与粗砂间来回悬停，拖拽又放回；视线在材料孔径属性表与设计画布间高频往返；自言自语。
		视觉：视线频繁跳跃，伴随微眯眼或皱眉。	
		姿态：身体前倾，出现迟疑性语气词。	

4.1.2. 介入策略：基于认知适宜性的三项原则

触发时机的确立仅是介入的第一步，智能体的介入策略还须接受认知适宜性的严格规制，以确保技术支持不破坏学习者的主体性。本研究提出以下三项原则：第一，认知克制。智能体在触发条件满足时，优先提供反思性提问或去语境化的部分线索；在算法设计上，采用信息屏蔽策略，通过约束指令强制大模型过滤掉直接答案或完整方案，以迫使学习者完成最后一步的闭环。第二，合意困难。智能体的支持要维持必要的认知难度，难度应动态调整，确保始终略高于学习者当前的舒适区。第三，必要时介入。智能体须抑制过度介入倾向，将试错机会保留给学习者。在算法机制上，通过强制系统在检测到停滞信息后延迟特定时长再触发响应，为学习者保留独立思考的缓冲空间。

4.2. 学习支架设计：多层次、最小化的跨学科学习支架

在情境触发机制中，智能体提供的干预资源必须精准匹配跨学科任务的复杂性。在此，有必要对多模态生成式支架与传统教学支架进行区分。传统的数字支架多为预设的、静态的、结构化的单向文本提示，缺乏对具体语境的感知；而本研究提出的多模态生成式支架，其核心优势在于能根据学生当前的具体错误步骤，实时生成结合了图解、语音和高亮线索的个性化提示。针对跨学科任务中多维度的认知障碍，本研究构建了多样式的支架，包含概念性支架、方法性支架、以及工具性与案例性支架。

为具象化这一干预过程，以跨学科任务“设计环保水净化工程装置”为例：当学生在“过滤材料分层选择”环节遭遇卡顿，传统系统可能只会因为答案错误而弹出文本提示“请检查物理过滤颗粒的大小顺序”。但在本框架下，多模态智能体通过数据融合发现：学生在当前界面的停留时间远超平均值，且频繁在两种粗细砂石间来回拖拽又取消，同时伴随困惑的微表情。综合判定学生陷入“策略犹豫”后，智能体主动触发干预：在屏幕侧边生成一个动态的水流过滤透视图，并配合柔和的语音引导：“观察图中水流穿过不同孔径材料时的流速变化，你觉得哪一层的缝隙能最先拦截大颗粒杂质？”这种紧贴当前操作断点、融合了物理现象与工程思维的生成式引导，不仅保留了学生最后一步的思考主权，更生动诠释了情境触发学习将情境从静态背景转化为动态解题线索的作用机理。

4.3. 高阶调节层：元认知与社会性支持

本层级旨在借助多模态智能体推动学习者从自我调节学习迈向社会调节学习。智能体在该层级中承担认知镜像的角色，在特定的认知出现偏离时被激活：

在元认知支持层面，针对盲目试错与思维跳跃等典型失效模式，智能体通过弹出式反思支架使学习者暂停当前操作，重审策略的有效性与证据的完备性。在社会性支持层面，面对个体无法消解的认知障碍，系统依据算法模型引入社会性资源支持，将个体困境转化为协作探究契机（如同伴互助或师生对话），通过社会性交互实现认知的协同修正。

贯穿上述支持的底层逻辑是预设认知失衡观察区间。在区间内，即便系统监测到困难，也会暂时抑制介入，允许学习者经历必要的认知挣扎与自主反思。只有当困难强度突破预设阈值，高阶支持机制才会被触发。上述必要时介入的策略设计，旨在防止技术支持对学习者的自我调节能力的过度替代。

4.4. 评价与迭代机制：过程分析与系统闭环

为突破传统评价对学习发生机理的遮蔽，本研究引入过程性数据分析，构建多模态评价机制。该机制致力于打开学习过程的“黑箱”，通过对高频行为的深度解析，建立从微观行为特征到宏观素养维度的映射关系。智能体针对跨学科学习的非线性特质，实施语义轨迹分析，具体涵盖三个维度：在时序维度，关注操作序列背后的认知转折点；在效能维度，量化支架触发后的行为修正率；在结构维度，引入认知网络分析，绘制问题解决的动态结构图。

上述分析使偏离标准路径但具有创造性的非典型解法得以被识别和肯定，最终生成包含深度理解、整合能力与策略灵活性的素养画像。

评价数据并未止步于反馈，而是回流至智能系统，构成强化学习的反馈回路。每一条后验数据都作为修正信号，驱动智能体动态修正对特定学习者的 ZPD 预测。该机制形成了“人机协同”的进化逻辑，系统随交互历史不断优化触发策略，实现技术支持与学习者发展的协同演进。

5. 研究结论与展望

5.1. 研究总结与理论贡献

针对跨学科情境化学习面临的“静态化”困境，本研究重新界定了智能技术支持下的学习者与情境的关系。研究表明，情境已从被动的学习容器转化为具备计算感知能力的主动交互体。在该框架中，多模态智能体超越了单纯的工具属性，转而承担认知协作者的功能，通过必要时介入和最小化生成式支架供给，在维持认知难度与提供支持之间达成平衡。基于过程性数据分析，系统获得自适应优化的能力，可随学习者能力的增长而动态校准触发阈值。

在学理层面，本研究致力于推动情境认知理论的可计算化转向。区别于传统研究侧重情境的社会学描述，本研究验证将模糊情境要素转化为可计算、可推断参数模型的可行性，为解析“情境-智能体-学习者”三元交互结构提供量化分析框架。在实践层面，本研究提出的必要时介入和合意困难原则，为缓解人工智能时代的认知外包风险提供具体的设计策略与约束机制。

5.2. 研究局限与未来展望

本研究构建的理论框架在理想条件下展现一定的可行性，但仍面临多重技术瓶颈与伦理挑战：在技术可行性方面，关键瓶颈集中在多模态信息的无感化采集和生成性输出的稳定性。真实学习环境噪音、光线复杂，仅靠普通摄像头与麦克风等设备无法精准捕获学习者的微表情、视线轨迹与交互动作，也很难在不干扰正常学习进程的同时完成全量数据的无感采集；生成内容的呈现形式多为文本、图片和视频，交互性不足，且生成内容存在输出不稳定和幻觉问题，需通过指令微调与人工校验进一步优化。在伦理风险防范方面，最大的挑战在于智能体过度介入所引发的认知外包和习得性无助。未来对这种风险的防范需要融入到系统的底层算法设计中，除了本研究提到的信息屏蔽和延迟干预外，还可以设计反向触发机制。该机制包括两个维度：一方面，当监测到学习者过于依赖智能体提示时，主动提升获取支架的难度；另一方面，当多模态数据监测到学习者认知过程过于顺畅或陷入机械化解题套路时，

系统应主动转换为提问者或搅局者，引入情境变量突变或抛出非典型视角进行质疑。这种主动制造合意困难的做法，不仅能防范认知惰性，更能推动学习者突破思维定势。

综上，未来研究需开展跨学段的纵向追踪，在不断优化多模态无感采集精度的同时，重点考察学习者在长期人机协同环境下的认知依赖性演变。通过技术精度与教育伦理的双向校验，真正实现智能技术对高阶思维发展的适度且良性赋能。

参考文献

- 于泽元、那明明 (2023)。情境化学习：内涵、价值及实施。华东师范大学学报（教育科学版），41(01)，89-97。
- 王萍、石磊和陈章进 (2018)。智能虚拟助手：一种新型学习支持系统的分析与设计。电化教育研究，39(02)，67-73。
- 卢宇、余京蕾、陈鹏鹤等 (2023)。生成式人工智能的教育应用与展望——以 ChatGPT 系统为例。中国远程教育，43(04)，24-31+51。
- 卢宇、余京蕾、陈鹏鹤等 (2023)。多模态大模型的教育应用研究与展望。电化教育研究，44(06)，38-44。
- 李吉林 (2016)。为儿童学习构建情境课程。中国教育学刊，(10)，4-7。
- 朱荣光 (2021)。触发深度学习的问题和情境。人民教育，(23)，65-67。
- 闫寒冰、杨淑婷、余淑珍等 (2025)。生成式人工智能赋能沉浸式学习：机理、模式与应用。电化教育研究，46(02)，64-71。
- 牟智佳、武法提 (2015)。电子书包中基于学习者模型的个性化学习资源推荐研究。电化教育研究，36(01)，69-76。
- 杨现民、余胜泉 (2015)。智慧教育体系架构与关键支撑技术。中国电化教育，(01)，77-84+130。
- 陈春芳 (2022)。核心素养导向下的情境化教学策略。中小学管理，(11)，39-41。
- 李嘉瑶 (2024)。基于智能技术的多模态学习分析：内涵、模型、挑战。中国教育技术装备，(01)，26-30+37。
- 郑金明 (2025)。基于 AI 自适应学习系统的个性化教育模式研究。大学，(05)，32-35。
- 许琼 (2024)。国外自适应学习系统的特点及启示。中国教育技术装备，(16)，142-147。
- 张家华、胡惠芝、黄昌勤 (2022)。多模态学习分析技术支持的学习评价研究。现代教育技术，32(09)，38-45。
- 张艳明、李巍巍、桂忠艳、孙尧 (2023)。基于混合学习模式的多模态学习分析机制探究。信息技术，(07)，6-63。https://doi.org/10.13274/j.cnki.hdzj.2023.07.011
- 教育部等七部门 (2025)。教育部等七部门关于加强中小学科技教育的意见。中华人民共和国教育部。http://www.moe.gov.cn/srcsite/A06/jcys_jyzb/202511/t20251111_1419878.html
- Clow, D.(2012). *The learning analytics cycle: Closing the loop effectively. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, 134–138.
https://dl.acm.org/doi/10.1145/2330601.2330636
- Dwivedi, Y. K., et al. (2023). *Opinion Paper: So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management*, 71.
https://doi.org/10.1016/j.ijinfomgt.2023.102642
- Grassini, S.(2023). *Shaping the future of education: exploring the potential and consequences of AI and ChatGPT in educational settings. Education Sciences*, 13(7), 692.
https://doi.org/10.3390/educsci13070692

- Hwang, G. J., Tsai, C. C., & Yang, S. J.(2008). *Criteria, strategies and research issues of context-aware ubiquitous learning. Educational Technology & Society*,11(2), 81-91.
<https://eric.ed.gov/?id=EJ814097>
- Paulo Blikstein, Marcelo Worsley.(2016). *Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. Journal of Learning Analytics*, 3(2), 220–238.
<http://dx.doi.org/10.18608/jla.2016.32.11>
- Sidney D’Mello, Blair Lehman, Reinhard Pekrun, Art Graesser.(2014). *Confusion can be beneficial for learning, Learning and Instruction, Volume 29, 2014, Pages 153-170.*
<https://doi.org/10.1016/j.learninstruc.2012.05.003>