

小学科学课堂教学 AI 评课系统： 深圳地区的探索与实践

童海云^{1*} 王亚琼² 辛海洋³

(1. 深圳市教育科学研究院, 深圳, 518081;

2. 深圳市福田区新沙小学, 深圳, 518048;

3. 香港中文大学, 深圳, 518172)

摘要: 评课是通过同行评议提升教师专业素养的重要方式和机制, 其核心价值在于推动平等交流、自主发展和提高教学质量。AI 评课技术的引入为课堂教学分析提供了新范式, 可以克服传统人工评课存在的效率低、主观性强、标准化不足等局限, 增加评价的维度, 提高反馈的敏捷性, 整合师生互动、参与度、情感状态等数据生成全景画像。本地区基于小学科学课堂教学实践, 开发了 AI 评课系统, 该系统的评课框架遵循还原性、多维度、可视化、敏捷性设计原则, 通过语音识别与大数据技术还原真实教学场景, 以活动光谱图、问题链树形分析等可视化图像对课堂教学进行分析诊断, 助力教师更客观地分析自己的教学特征和问题, 助力科学课堂实现从“知识传输”到“素养发展”的转型, 提升教师专业素养。

关键词: 教师专业素养; 课堂评价; AI 评课系统; 教学活动光谱图; 问题树形分析

评课作为教学研究的重要活动, 对教师专业发展具有不可替代的支撑作用。评课一般是经历“课堂观察指标设定—课堂教学活动观察—观察后的判断与分析—交流和讨论—形成诊断性报告”等几个主要过程, 通过系统性评估揭示教学行为的显性特征, 其核心在于“从教育理论高度对课堂行为作出解释”^①, 通过数据指标量化教学效果。而在教学实践中, 评课通常聚焦于“交流和讨论”的评议活动, 从“以学论教”的角度评价教学, 通过研究学生的学习效果反思教师的教学行为, 构建起教师反思的认知闭环, 是教师教学素养提升的重要路径。

* 童海云, 51887538@qq.com。

① 卢立涛, 余雨芳, 翁雨轩. 循证视域下小学数学听评课教研活动实施的个案研究 [J]. 北京教育学院学报, 2025, (01): 54.

显然，课堂观察指标设定是评课活动的发起点和评议结果的依归，指标设定尤其重要又充满创造性，对教学诊断性报告和教学行为改进起到指南针作用。指标设定有通用型和学科专属型之分。LICC（Learner, Instruction, Curriculum, Classroom Culture，即学生学习、教师教学、课程性质、课堂文化）范式四维度就是通用型的代表，主要从学生学习、教师教学、课程性质、课堂文化四个维度解构课堂生态，覆盖了教学的核心要素，能够促使教师从教材解读、教学设计到课堂呈现等多方面进行深入思考，从而提升教学质量^{①②}。学科专属型通常是针对某一学科，根据其学科课程标准、学科教学内容和方式的特点、教师专业背景知识等设定一些特定的指标，以更好地满足学科教学所需，还可以具体通过教材文本解读、教学设计与实施等维度的分析，形成教学行为的系统化反思路径^③。学科专属型还可以根据地区学科教学发展所需进行指标定制。我们这项研究是基于深圳地区的小学科学课堂教学实践经验，开发了面向小学科学教学素养培养的学科专属型人工智能评课系统。本文将对此人工智能评课系统的设计原则和框架进行介绍，并举以实例进行解读，旨在与同行进行交流，抛砖引玉，完善我们的工作。

一、人工评课的实践优势与时代困境

人工评课的实践优势主要体现为：1. 情境化解读能力：专家型教师对课堂生成性问题的敏锐捕捉，比如，课堂教学过程中教师打断行为的意图分析，学生非言语反馈的即时捕捉等^④。2. 教育智慧的不可替代性和传承性：科学探究活动的情感温度评估，师徒制评课中的隐性知识传递^⑤。3. 个性化反馈：传统人工评课能够根据学生的具体需求提供个性化的反馈和建议，教师熟悉学生的特点，可以针对其弱点提供定制化的指导^⑥。4. 过程评估：注重学生的学习过程，包括学习态度、参与度和学习方法等，而不仅仅是结果评估^⑦。

人工评课的实践困境主要表现为：1. 评价标准的模糊性：人工评课容易受到教师主观判断的影响，导致评价结果的不一致性和偏差，主观经验与客观指标的失衡^⑧还会导致多人评课时“公说公有理、婆说婆有理”的窘境。2. 数据采集的局限性：人工进行课堂观察时会因观测维度的不同出现“盲人摸象”现象。3. 反馈时效的滞后性：人工评课需要耗费大量时间和精力，特别是在面对大规模课堂教学时，难以实现即时反馈，造成教学改进的“热反馈”缺失^⑨。4. 复杂能力的评估困境：对

① 吴克虎，杨志惠. 构建观课议课模式促进教师专业成长 [J]. 教育文汇，2019，(07)：11.

② 董文萍. 观课议课对青年教师成长的作用 [J]. 课程教育研究，2012，(20)：169.

③ 寇永升. “评”出一堂语文课 [J]. 语文教学通讯，2022，(16)：41.

④ 同③.

⑤ 张定强，苏丹兰. 专业发展视域下教师评课：理念责任、知识能力、实践体验 [J]. 教师教育论坛，2013，(09)：17.

⑥ Kakkonen T, Sutinen E. Semi-automatic Assessment Model of Student Texts: Pedagogical Foundations [J]. ArXiv, 2012, (02)：183.

⑦ Du H. Using Artificial Intelligence Technology to Improve Interaction and Learning Assessment in Higher Vocational English Classes [J]. The Educational Review, USA, 2023, (12)：1920.

⑧ Wang G. Research on the Construction of Classroom Teaching Evaluation System Based on Artificial Intelligence [J]. 2021 International Conference on Big Data Engineering and Education (BDEE), 2021, (01)：101.

⑨ Liang Y, Wu J. Research on the Method and Path of Embedding Artificial Intelligence into Higher Vocational Teaching Reform [J]. Frontiers in Educational Research, 2023, (08)：53.

学生创造力、批判性思维等能力的评估存在困难,对学习能力、创新思维等维度缺乏有效观测工具^①。

二、AI 评课的优势

在新的时代背景下,以数据驱动的课堂观察和诊断为主要特征的人工智能评课(AI 评课)呈现出星火燎原之势,正在被越来越多的学校、教育行政部门和教师所采用,成为教学数字化和数智化的革新性应用。

相对于传统的人工评课,AI 智能评课的革新性主要体现在:

1. 数字化赋能。数字化赋能可以兼顾规模化和个性化教育对评价效率的需求^②。规模化教育是指通过标准化、系统化和数字化的方式,实现教育内容、教学资源和学习过程的大规模生产和传播,以满足广泛学习者的需求。规模化教学也需要同时兼顾以学生自身成才、潜能发展、创新人才培养为目标的个性化学习需求。AI 评课可以提供精准化的数据支撑^③:通过分析学生问题链响应差异,AI 评课可以推荐分层学习任务(如低认知水平学生侧重“现象观察”,高水平学生设计对比实验),帮助教师实现因材施教。此外,AI 课堂观察工具可辅助教师设计符合国际学生能力评估计划(Programme for International Student Assessment,简称 PISA)的探究活动,缩小城乡科学教育差距,推动教育优质均衡发展。

2. 评课的技术路径革新。主要表现为多模态数据融合,视频分析、语音识别、文本挖掘的协同应用^{④⑤⑥};教学行为建模,基于深度学习的教师行为特征提取^⑦,生成可视化报告(如问题链分析图、专注力热力图),帮助教师快速定位低效提问或认知超载环节,推动其从“经验型”向“研究型”转型,促进教师专业发展;专家系统的知识沉淀,将名师经验转化为评价算法。

3. 范式革新。主要体现在:全流程覆盖,从单次课例分析到教学周期跟踪^⑧;多维度透视,融合“师—生”行为分析、认知负荷测算等创新指标^⑨,精准诊断教学短板;动态评价标准的生成,基于教学大数据的指标优化^⑩、地区发展和教师成长进阶所需,动态地优化和建构评价体系。

4. 数字共同体建构。建立 AI 评课的教学案例库,更方便将隐性经验通过网络转化为可交流的显性知识^⑪,通过教学共同体的构建以评课为纽带形成教师发展共同体^{⑫⑬}。通过多维反馈促进

① Meylani R. A Comparative Analysis of Traditional and Modern Approaches to Assessment and Evaluation in Education [J]. Batu Anadolu Eğitim Bilimleri Dergisi, 2024, (04): 520.

② 于海平,孙萍,彭玲.新时代背景下高等职业教育评价改革与评价体系研究 [J].现代商贸工业,2024,(21):230.

③ 宗锦莲.深度学习理论观照下的课堂转向:结构与路径 [J].教育学报,2021,(01):59.

④ 孙众,吕恺悦,骆力明,等.基于人工智能的课堂教学分析 [J].中国电化教育,2020,(10):15.

⑤ 刘清堂,何皓怡,吴林静,等.基于人工智能的课堂教学行为分析方法及其应用 [J].中国电化教育,2019,(09):13.

⑥ 赵刚,朱文娟,胡碧灵,等.多维度特征融合的教学行为智能分析模式构建 [J].电化教育研究,2020,(10):36.

⑦ 吴冬丽.基于视频的教师行为识别研究 [D].华中师范大学,2021.

⑧ 李玲,肖君,顾小清.智能课堂分析的多维指标体系构建:面向开放教育 OMO 智慧教学环境的研究 [J].远程教育杂志,2023,(01):48.

⑨ 刘清堂,何皓怡,吴林静,等.基于人工智能的课堂教学行为分析方法及其应用 [J].中国电化教育,2019,(09):13.

⑩ 吴立宝,曹雅楠,曹一鸣.人工智能赋能课堂教学评价改革与技术实现的框架构建 [J].中国电化教育,2021,(05):94.

⑪ 孙众,吕恺悦,骆力明,等.基于人工智能的课堂教学分析 [J].中国电化教育,2020,(10):15.

⑫ 张定强,苏丹.专业发展视域下教师评课:理念责任、知识能力、实践体验 [J].教师教育论坛,2013,(09):17.

⑬ 张茂宇.基础教育师资培训中应注重教师观课议课能力的开发和培养 [J].新课程(教研),2010,(08):158.

个性化教学风格的形成，指引教师不断尝试新的教学策略，推动课堂教学改革通过评课数据发现教改突破点^①。

三、本地区 AI 评课系统的设计特点

基于上述分析，结合众多 AI 评课系统的架构、学科特征、学段特点、课程文化和价值、地区发展等要素，我们设计了小学科学课堂教学 AI 评课系统——突出学科教学素养为主的混合型评课系统，主要表现为四维架构设计特点。

（一）还原性

还原性原则源于哲学中的还原论，主张将复杂系统分解为独立组件进行分析与重构，通过“整体等于部分之和”的逻辑实现系统功能。在教育场景中，其核心在于通过分解教学行为、知识结构、师生互动等要素，精准还原课堂动态教学发生过程。

（二）透析性

借助专家型知识沉淀，设计多个维度指标的关联性和纵深度观察，结合学生的核心素养发展的需要，实现多维度透析课堂。

1. 问题树形分析：追踪教师提问的认知层级（如布鲁姆分类法），评估学生从“记忆事实”到“创造性解决”的思维跃迁。
2. 行为与情感监测：利用词频分析、对话记录等技术，量化学生参与度、好奇心及环境责任感。
3. 数字学习能力评估：通过学生在数字工具中的操作路径，分析其自我调节策略与计算思维水平。

（三）可视化

在报告的呈现形式上进行创新，用众多可视化的图表来呈现分析结果，让数据可视，有助于教师一目了然掌握教学行为动态变化，精准分析教学过程，客观呈现教学问题。

（四）敏捷化

短时反馈评课结果，提高反馈热度，变冗长的报告文本为模块化标注，有助于教师第一时间了解教学问题，进行教学设计及方法改进。

四、AI 评课的地区性实践与探索

经过系统性开发和技术上优化后，我们在深圳地区进行了实践与探索，在实践中不断优化指标体系，在探索中增强学科属性，在周期性迭代中增强教师友好性，使 AI 评课更加好用、适用。下面介绍评课系统的部分实践探索。

^① 王绥联，引领课程教学之钥——谈校长如何观课议课 [J]，新教育，2020，（22）：10.

（一）教学活动光谱图

1. 设计原则

教学活动光谱图是一种将课堂教学活动按照教师教学指导、学生自主性和实践复杂性进行分层的可视化工具，旨在帮助教师设计、实施与评估科学课堂中的探究活动。它通过明确不同层级的探究目标与行为特征，为科学教学提供结构化支持。光谱图可以有效聚焦于探究实践活动，在以学生为主体的课堂，观察课堂时间分配、学生是否主动学习以及活动之间是否有连续性和结构性。通过可视化呈现，教师可有针对性了解自己在课堂过程中的时间分配。

光谱图可以应用到单课教学、同课多次、大单元教学连续性、不同年段课例对比、不同领域课例对比等教学评估。不同的应用场景可以根据光谱的具体分布情况进行评估，从而提出课堂教学的诊断性建议。

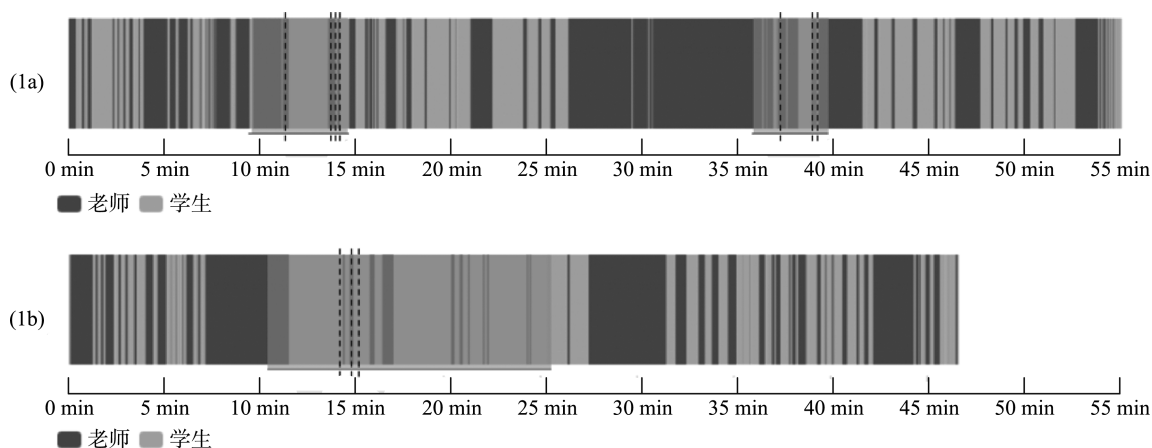


图1 《地球——宇宙的奇迹》前后两次教学活动光谱图

2. 应用举例

我们选取同一教师连续执教相同教学课题的两节高年段的科学课，将其进行对比分析。光谱图采用双色编码系统，深色区块表征教师主导时间，浅色区块表征学生自主时间，时间轴精确记录教学环节的时序分布，分块区域代表集中的学生探究活动，虚线表示教师强制中断教学活动的教学行为。

首节课例显示（图1a），总课时55分钟的课程呈现典型的交替式师生互动结构。教师主导行为占比较高，集中分布于课堂中段（第23—30分钟）的连续讲授环节。学生自主活动设置两次，分别在第10分钟、第35分钟开展，但单次时长均未超过5分钟。光谱中的虚线表示教师有2次强制中断的教学行为。光谱图中学生和教师行为高频次的交互和教师的中断行为都表示这节课学生探究进程是比较碎片化。

上课教师根据光谱图的诊断报告，对同一课题进行了教学改进和实施。通过对比分析可以发现，第二节课例（图1b）教学活动的结构有了显著的优化。课时严格控制在46分钟，其中集中式学生自主探究时段延长至15分钟，符合高段教学自主化探究层级的时长建议。教师强行中断学生学习活动的频次下降为1次。光谱分布结构分析显示，该课例的探究连续性有所提升，师生互动模式从交替式转向模块化。

对比研究发现,两节课例在自主探究的时间分布上呈现显著差异:首节课采用分散式短时探究(2次<5分钟),次节课实施集中式长时探究(15分钟)。这种转变印证了探究活动连续性指标对教学活动的提示作用。此外,教师介入频次有所下降表明光谱图可视化反馈对教师行为改进的促进作用。

3. 诊断与指导

A. 指引科学探究的连续性设计

光谱图可以根据年段将探究活动形象化表征,通过图谱指引教师在教学中构建科学探究“低自主”到“高自主”的探究连续体,即验证性探究(教师主导)、引导性探究(半开放)和开放性探究(学生自主)三个层级。

B. 优化教学时效

科学课采用探究实践为核心教学方式,结构性更强的教学活动才有可能让探究实践真正落地。光谱图的分布结构帮助教师观察和诊断教学中学习内容与时间分配的合理性,也能更好地指引教师优化教学时效,进行探究活动结构性设计。

C. 关联学生专注度和参与度

学生探究活动时长根据实验类型划分,同时结合负面(如分心、被动等待)占比,评估探究活动的吸引力与适切性,以评估学生学习活动的专注度。同时,光谱图可以通过记录学生主动提问、协作讨论、实验操作等行为的频次与时长,观测教学活动中学生的参与度。

(二) 问题链树形分析

1. 设计原则

A. 问题解决关联观念发展

科学观念的发展是需要事实、逻辑和反省抽象的多重建构。科学教学中通常通过问题解决的方式来实现科学观念建构。大问题需要分解成学生能够解决的小问题,通过小问题去获得事实支撑,以“大问题—小问题—小事实—小逻辑—小观念—大观念”的进阶性理解,为学生构建科学观念发展奠定基础。课堂问题透析将观察和评估学生科学观念发展的支撑性逻辑。

2. 问题链逻辑反映科学思维和社会建构

在解决具体问题时,需要教师进行具体问题的问题链设计。问题链设计有两个方面的逻辑需要兼顾:一是科学思维逻辑,可以通过模拟科学家“情境—提出问题—设计实验—验证结论”的思维路径,契合科学教育中“科学本质观”(Nature of Science, NOS)的核心目标。二是知识的社会性建构逻辑,可以通过“问题—反馈—评估—扩展”的知识对话模式,通过社会性知识建构的方式使课堂各要素(教师提问、学生回应、知识逻辑)进行动态交互,体现“以学习者为中心”主动建构知识的教学框架。

C. 树形分析

通过问题的树形呈现和分析,可以准确抓住“驱动性问题”的源头,可以根据问题流向的分支和总需关系分析对教学活动的意义逻辑进行整体性评估,可以结合问题的源头和终点分析对教学目标的实现进行评估。

2. 应用举例

下文以《地球——宇宙的奇迹》的第二次上课为例进行问题链树形分析。

先对教师课堂上提出的问题进行还原性提取和分类处理。本课教师共提出了34个问题,见表1。

表 1 《地球——宇宙的奇迹》课堂教学教师提问提取与分类处理表（共 34 个问题）

编号	原始序号	简化后的提问内容	开始时间	布鲁姆分类
1	12	视频中提到火星曾经可能具备什么条件？	00：00：47	理解型（信息提取）
2	14	你认为火星能否成为人类第二个家园？为什么？	00：00：58	评价型（价值判断）
3	36	为什么多数同学认为火星可能宜居？	00：02：06	分析型（归因推理）
4	60	选择“不确定”的同学认为火星的哪些条件不明确？	00：03：29	分析型（要素拆解）
5	73	判断星球是否宜居的标准是什么？	00：04：18	综合型（标准建构）
6	79	地球的生存条件与其哪些本质特征相关？	00：04：39	分析型（关联分析）
7	82	地球为生命提供了哪些基础生存条件？	00：04：57	知识型（事实回忆）
8	84	地球的氧气和液态水如何形成？	00：05：03	应用型（机制解释）
9	96	地球与太阳的距离如何影响温度？	00：05：32	理解型（因果解释）
10	111	生存还需要哪些未被提及的条件？	00：06：17	综合型（要素补充）
11	122	地球的栖息地特征包含哪些要素？	00：06：41	分析型（特征归纳）
12	137	地球的哪些特征支持生命所需条件？	00：07：18	综合型（系统分析）
13	139	适宜温度的形成机制与哪些因素相关？	00：07：22	分析型（机制探究）
14	141	其他生存条件与地球的哪些特点有关？	00：07：34	应用型（关联迁移）
15	176	地球的特征如何影响生命条件？	00：09：25	综合型（系统关联）
16	182	虚拟实验能揭示地球特征的哪些关键信息？	00：09：38	应用型（实验设计）
17	214	日地距离改变如何影响温度变化？	00：11：12	应用型（变量控制）
18	255	设备识别问题是否影响实验结论可靠性？	00：14：06	评价型（误差评估）
19	358	地球质量变化如何影响大气层？	00：24：17	应用型（模拟验证）
20	366	地球质量极端变化会产生什么后果？	00：24：33	综合型（推理论证）
21	460	如何系统呈现各组探究成果？	00：30：29	综合型（成果整合）
22	471	地球磁场对生物活动有何具体影响？	00：30：58	应用型（功能解释）
23	519	磁场对地球生态系统的保护作用如何实现？	00：33：16	分析型（机制剖析）
24	532	地球自转周期如何维持生命节律？	00：33：57	理解型（系统关联）
25	599	引力变化如何影响大气层稳定性？	00：37：17	应用型（变量关联）
26	608	大气厚度如何调节昼夜温差？	00：37：37	分析型（原理阐释）
27	633	火星当前存在液态水的证据是否充分？	00：38：45	评价型（证据评估）
28	640	火星大气成分是否满足基本生存需求？	00：39：00	分析型（要素对比）
29	665	火星土壤特性对植物生长有何限制？	00：40：13	应用型（条件分析）
30	682	火星环境改造的可能性有多大？	00：41：09	综合型（可行性评估）
31	732	地球条件的协同作用如何形成生命基础？	00：44：08	综合型（系统思维）
32	738	宇宙中类地行星存在的概率有多低？	00：44：23	评价型（概率认知）
33	751	地球的独特性对人类文明发展有何启示？	00：44：52	评价型（价值判断）
34	783	如何将地球保护意识转化为具体行动？	00：46：30	应用型（实践迁移）

其次是问题链解构与分组。此环节是在课堂提问还原性提取的基础（表 1）上，结合布鲁姆分

类法等工具标注认知层级。随后，依据问题间的逻辑关系进行解构和分组，例如识别“核心大问题”与围绕其展开的“条件解构”“实验验证”等子问题群。这种分组模式学术上反映了科学教学中“大问题—小问题—…—大观念”的进阶性理解路径，并模拟了科学思维逻辑（如情境—问题—实验—验证，契合科学本质观）及知识的社会性建构逻辑。此分析评估了教师提问能否系统地引导学生认知建构。

其次是逻辑关系特征捕获。这部分深入剖析问题链整体结构的学术意义。报告捕获了如树状渗透结构、双重验证机制、镜像对比论证和认知梯度递进等特征。这些特征揭示了教学设计的逻辑严谨性与认知引导策略：双重验证体现科学认识论中理论与实践的结合；认知梯度递进评估提问能否有效引导学生进行高阶思维，从事实认知逐步升华至价值判断和实践应用。这部分分析评估了提问结构如何服务于教学目标和素养培养。

最后是问题关系树形图。这是一个重要的可视化工具。它将抽象的问题链逻辑关系与分组结构以树状图形式直观呈现。图中运用实线箭头表示强逻辑递进关系，虚线箭头表示数据支撑或对比关系，分支结构表示论证维度分化，配合字母编号代表不同问题组或维度。这种可视化符合 AI 系统“可视化”原则，有助于教师清晰理解提问设计的整体结构、分支与关联，从宏观层面诊断教学逻辑、评估教学目标实现情况。

相关的分析过程和 AI 输出的分析报告见图 2。

我们将表 1 中的问题和图 2 中的树形分析结合起来，就可以看到整节课的问题框架和认知逻辑：

A 核心问题：判断星球是否宜居的标准是什么？

B 支链主要问题：Q7 地球为生命提供了哪些基础生存条件？

Q9 地球与太阳的距离如何影响温度？

Q12 地球的哪些特征支持生命所需条件？

Q14 其他生存条件与地球的哪些特点有关？

B 支链所有问题在于引导学生分析“地球生命所需的基本条件”，通过分析学生用地球的基本条件作为标准来判断其他星球是否宜居。

C 支链主要问题：Q16 虚拟实验能揭示地球特征的哪些关键信息？

Q17 日地距离改变如何影响温度变化？

Q19 地球质量变化如何影响大气层？

Q26 大气厚度如何调节昼夜温差？

这组问题是学生做分组数字虚拟实验的引导性问题。学生做完虚拟实验后，借助支链问题分析影响地球表面温度的相关因素：日地距离、地球质量、大气厚度等。通过分析，学生能够理解地球处于宜居带、地球自身引力和大气层影响着地球表面温度、其他星球应当具备类似条件才有可能宜居。

D 支链主要问题：Q2 你认为火星能否成为人类第二个家园？为什么？

Q27 火星当前存在液态水的证据是否充分？

Q28 火星大气成分是否满足基本生存需求？

Q29 火星土壤特性对植物生长有何限制？

学生借助教师提供的资料和 AI 对话，分析火星是否适宜人类生存。由于前面 B、C 支链问题

分析报告：《地球——宇宙的奇迹》问题树型分析

一、问题链解构与分组

1. 核心大问题

Q0 (星球宜居性): 如何判断星球是否宜居? 地球的宜居条件具有何种独特性? (对应问题 5、7、15)

2. 紧密跟随的子问题群

(1) 条件解构组:

Q5 (标准建构) → Q7 (基础条件) → Q9 (距离影响) → Q12 (系统关联) → Q14 (补充要素)

(验证路径: 从条件识别到系统分析)

(2) 实验验证组:

Q16 (实验设计) → Q17 (变量控制) → Q19 (模拟验证) → Q26 (原理阐释)

(科学方法: 通过虚拟实验验证理论假设)

(3) 星际应用组:

Q2 (火星评估) → Q27 (证据评估) → Q28 (要素对比) → Q30 (可行性分析)

(知识迁移: 将地球标准应用于外星评估)

(4) 认知深化组:

Q32 (概率认知) → Q33 (价值判断) → Q34 (实践迁移)

(认知升华: 从科学事实到文明责任)

二、逻辑关系特征捕获

• 树状渗透结构: 核心问题 (Q0) 同时向四个论证维度延伸, 形成“标准-验证-应用-认知”的立体框架

• 双重验证机制: 理论标准 (条件组) 与实验验证 (实验组) 形成闭环论证

• 镜像对比论证: 地球条件分析 (条件组) 与火星缺陷分析 (应用组) 构成对比验证

• 认知梯度递进: 从事实认知 (Q7) 到概率认知 (Q32), 最终导向价值判断 (Q33) 和实践要求 (Q34)

三、问题关系树型图

A (核心: 宜居标准)

├ B (条件解构)

│ │ B1 (标准建立-Q5) → B2 (基础条件-Q7) → B3 (距离影响-Q9)

│ │ B4 (系统关联-Q12) → B5 (补充要素-Q14)

├ C (实验验证)

│ │ C1 (实验设计-Q16) → C2 (变量控制-Q17)

│ │ C3 (模拟验证-Q19) → C4 (原理阐释-Q26)

├ D (星际应用)

│ │ D1 (火星评估-Q2) → D2 (证据评估-Q27)

│ │ D3 (要素对比-Q28) → D4 (土壤限制-Q29)

│ │ D5 (可行性分析-Q30)

├ E (认知深化)

│ E1 (概率认知-Q32)

│ E2 (价值判断-Q33)

│ E3 (实践迁移-Q34)

四、逻辑符号说明

• 实线箭头 (→): 强逻辑递进关系

• 虚线箭头 (⇨): 数据支撑或对比关系

• 分支结构 (├/└): 论证维度分化

• 字母编号: A (核心问题)、B(标准建构)、C(实验验证)、D (标准应用)、E(认知升华)

图2 AI的分析报告摘录:《地球——宇宙的奇迹》问题链树形分析

研究了地球适宜生命生存的条件, D支链问题就围绕着这些条件对火星进行宜居性评估。Q27和Q28两个问题比较合适, 但Q29问题明显偏离了原来研究的主线, 引导学生去研究土壤成分, 这是一个需要调整的问题。

借助问题链的树形分析, 我们可以清晰地判断出教师的教学活动设计的认知逻辑: 分析地球的生存条件—模拟研究地球的基本特征—评估火星是否宜居—总结地球的独特性—认识地球是宇宙的奇迹。树形分析还能评估支链之间关联程度, 如火星宜居性评估必须以地球的特点为依据;

还能进行个别问题的偏离性分析。

3. 诊断与指导

A. 提升教师的提问能力

教师需要更加关注问题链的逻辑解构与层级映射，主要体现为下面三个层级：

- 锚定大问题或者核心问题。以课程标准中的核心概念为锚点，提取大问题，确保问题链指向高阶素养目标。

- 网络化分析子问题。基于问题驱动理论（Question-Driven Learning, QDL），将大问题拆解为递进性子问题，采用逻辑网络图量化子问题与核心问题的关联强度。

- 增强问题价值判别。问题价值主要体现为三个方面——目标导向性（直指大问题解决）、认知挑战性（符合最近发展区）、生成启发性（引发新问题或迁移）。

B. 提升课堂学习活动的底层逻辑

问题链的动态分析认知路径分析，可以帮助教师更好地分析和设计课堂教学活动，突出其活动的认知导航价值。比如，通过问题树的分析，教学中教师更加关注“问题—事实—观念—迁移”四阶认知模型。问题驱动，大问题激活认知冲突；事实提取，通过子问题引导观察与数据收集；观念生成，整合事实形成解释模型；迁移验证，应用模型解决新情境问题。

C. 强化核心素养的系统性设计

在前面的案例分析中，我们可以明显观测到教学实施过程中科学观念建构的系统性设计：首先以标准建立（Q5 标准建立）为认知锚点，引导学生逐步解构地球宜居条件（如 Q7 地球温度基础条件、Q9 日地距离影响）；继而通过虚拟实验设计（Q16 实验设计）和变量控制（Q17 变量控制）实现理论验证，促进结构化认知的形成。问题设计充分融合布鲁姆分类目标，既包含具体运算阶段的事实性积累（如 Q7 地球基础条件回顾），又涵盖形式运算阶段的抽象推理（如 Q30 火星改造可行性分析），实现了认知水平的梯度提升。

AI 分析可以实现学习路径的偏离诊断，当子问题偏离大问题逻辑主线（如过度聚焦非核心细节）或跳跃层级（如未完成事实积累即要求迁移），AI 生成警示信号（如“逻辑断裂：缺失‘能量来源’子问题”），有助于教师调整教学活动和教学行为。

课堂问题链树形分析通过“逻辑解构—动态追踪—可视化诊断”三位一体模型，破解传统提问的盲目性与浅表化困境。其核心价值在于将问题链从“教学工具”升维为“素养生成载体”，助力科学课堂实现从“知识传输”到“素养建构”的范式转型。

五、结论与展望

通过地区的教学实践，我们设计和探索了混合型小学科学课堂教学 AI 评课系统，突出学科的核心素养导向。评课系统比较好地帮助教师复盘课堂教学，对教学活动、教学设问、建设性对话、探究实践专注度等进行多维还原，让教师能更客观、全面、细致、深刻地分析和反思教学问题，从而更好地提升教师的教学素养和学生的科学素养。

AI 评课系统依然面临诸多挑战，如多模态数据融合度还待提升，专家型的深度指导难以转化为评价算法，学生创造力、批判性思维等能力的观测工具效度较差，教师教学风格的个性化观测还难以实现。相信在较长的时间内，AI 评课和人工评课还会赋能共生，取长补短，相互发展。

作者简介

童海云，高级教师，深圳市教育科学研究院科学教研员，研究方向：科学教学。

王亚琼，高级教师，深圳市福田区新沙小学，研究方向：STEM 教学、科学教育、化学教育。

辛海洋，香港中文大学信息工程学博士，研究方向：教育数字化、STEM 教育。

AI-Based Classroom Evaluation System for Primary School Science Teaching: Exploration and Practice in Shenzhen Region

Tong Haiyun¹, Wang Yaqiong², Xin Haiyang³

(1. Shenzhen Institute of Education Science, ShenZhen, 518081;

2. XinSha Primary School, Futian District, ShenZhen, 518048;

3. The Chinese University of HongKong, ShenZhen, 518172)

Abstract: Teaching evaluation serves as a crucial method and mechanism for enhancing teachers' professional competence through peer review, with its core value lying in promoting equal exchange, autonomous development, and improving teaching quality. The introduction of AI technology offers a new paradigm for classroom teaching analysis, overcoming limitations inherent in traditional manual evaluations such as low efficiency, strong subjectivity, and insufficient standardization. It expands evaluation dimensions, enhances feedback agility, and integrates data on teacher-student interactions, engagement levels, and emotional states to generate comprehensive profiles. Based on primary school science classroom teaching practices in the region, an AI teaching evaluation system has been developed. The system's evaluation framework adheres to design principles of reducibility, multidimensionality, visualization, and agility. By employing speech recognition and big data technology to reconstruct authentic teaching scenarios, it utilizes visual representations such as activity spectrum graphs and problem tree analysis to diagnose classroom instruction. This assists teachers in more objectively analyzing their instructional characteristics and issues, facilitating the transformation of science classrooms from “knowledge transmission” to “literacy development,” thereby elevating teachers' professional competence.

Key Words: Teacher professional competence; Teaching evaluation; AI-based teaching evaluation system; Instructional activity spectrum; Problem tree analysis