



# AI 时代学评融合的概念、内涵与应用模式\*

张生<sup>1</sup>, 王雪<sup>1</sup>, 辛涛<sup>1①</sup>, 梁凯丽<sup>2</sup>

1.北京师范大学 中国基础教育质量监测协同创新中心, 北京 100875

2.北京市西城区教育科学研究院, 北京 100035

**摘要:** AI 时代对教育评价提出了新的需求, 传统以知识技能为核心、侧重甄别与筛选的教育评价体系, 已难以适应发展学生核心素养的战略需求。本文提出并系统阐述了“学评融合”这一新型评价理念与实践框架。研究首先界定了学评融合的核心内涵, 即将学习者被动参与的评价任务转换为以智能体赋能的学习者主动参与的评价学习活动, 借助智能技术形成一个“评价蕴含学习, 学习数据赋能评价, 评价结果驱动学习迭代”的闭环; 进而, 从“学”“评”“融合”三个视角解析了其内在结构、特征。研究进一步揭示了学评融合区别于传统评价理念的五个根本特性。在此基础上, 本文明晰了学评融合在个性化学习与高阶思维测评中的双轨应用模式, 并最终提出了以素养导向、文化重塑、系统联动与专业支撑为核心的四大实践策略。学评融合理念不仅为破解当前教育评价改革困境提供了理论支点, 更通过其系统化的实践路径, 为构建赋能学生全面发展、契合中国教育语境的新型评价生态指明了前瞻方向。

**关键词:** 学习评价; 人工智能时代; 学评融合; 评价改革

**中图分类号:** G434

**文献标识码:** A

\* 本文系 2021 年度国家重点研发计划“社会治理与智慧社会科技支撑”重点专项项目“大规模学生跨学段成长跟踪研究, 学生综合素质测评模型构建”(项目编号: 2021YFC3340801)研究成果。

① 辛涛为本文通讯作者。

## 一、引言

人工智能时代的到来与“新质生产力”等国家战略的提出<sup>[1]</sup>, 正以前所未有的深度与广度重塑社会结构与人才需求。2025 年 1 月, 中共中央、国务院印发了《教育强国建设规划纲要(2024—2035 年)》<sup>[2]</sup>, 纲要中指出为了支撑国家现代化建设和民族复兴, 亟需深化综合改革, 激发教育活力, 培养时代新人。在此背景下, 教育评价作为教育改革的“指挥棒”, 其改革方向与成效直接关系到教育强国建设的全局。传统的知识技能导向与“一刀切”式的评价理念, 已无法适应新时代对高素质、复合型人才的培养和选拔要求。因此, 依据《深化新时代教育评价改革总体方案》<sup>[3]</sup>等政策精神, 破除“五唯”痼疾, 建立科学、多元、动态的教育评价体系, 已成为教育领域的核心议题。

学术界与实践领域已围绕教育评价的转型展开了积极探索。一方面, 教育评价理论正在经历根本性重塑, 研究者们从哲学层面的本体论、认识论、价值论与方法论等多个维度进行了深入的理论思辨<sup>[4]</sup>。另一方面, 在技术变革的推动下, 教育评价实践也在向学习化、个性化、精准化和高效化加速转型。生成式人工智能、智能体(AI Agents)、多模态学习分析等技术的融合应用, 正推动评价体系从“工具辅助”走向“智能内生”。智能技术作为新型“数字底座”,

已能深度融入从内容生成、学习支架、过程互动到思维引领的完整链条, 为实现以学生自主探究与互动为核心的新型评价范式提供了坚实支撑<sup>[5]</sup>。然而, 在当前教育评价转型的初期阶段, 一个核心矛盾日益凸显: 评价的诊断性功能与学习性功能常处于彼此分离的状态。诊断结果往往难以有效转化为指导日常学习的精准策略, 而学习过程中的关键数据与表现也未能被评价体系实时捕捉与运用, 导致评价负担激增, 评价改进与发展作用大打折扣。如何使评价的诊断与学习两大功能在理论与实践达到相互融合、有机共生, 并在日常教学与测评考试中得以高质量的落地实践, 成为亟待解决的关键问题。为系统性地应对这一挑战, 本文提出以智能技术为内在驱动的“学评融合”理论与实践应用框架。本研究将深入阐释“学评融合”的内涵、独特性、应用场景与实践策略, 以期为新时期教育评价体系的深化改革提供有价值的理论参考与实践指引。

## 二、学评融合的概念、内涵、特征

评价兼具诊断学习成果与促进深度学习的双重功能。其功能重心随时代演变而不同, 在生成式人工智能与智能体等技术驱动下, 对传统认知技能的自动化、精准化、智能化、个性化诊断已日趋成熟, 这使得评价的重心得以向更深层次的功能转移。一方面, 评价的诊断性功能得以拓展,





程中的学习”。这指的是学生在亲身参与评价活动(如制定标准、审视作品、进行反思)时,所直接引发的高阶思维发展。这种学习是内嵌于评价行动之中的、即时性的成长。需要明确的是,这区别于“基于评价结果反馈后所进行的学习”——后者本质上是评价的导向与调节功能在发挥作用,其学习发生在评价过程之后。

在此框架外,教育中当然还存在一般意义的学习,例如自主探究、小组协作等。这类学习活动与评价的衔接关系,传统上属于“教学评一致性”的研究范畴。然而,学评融合范式并不排斥一般意义的学习,反而将其视为重要的数据源泉。在测量某些高阶能力(如计算思维)时,范式要求将一般学习过程设计为可观测、可评价的活动任务,使其产生的数据能够赋能于评价的诊断性功能。总而言之,学评融合的“学”,其本质是评价任务本身所蕴含的、即时的学习价值。而如何将“一般意义的学习”有机地融入“评价中的学习”与“评价中的诊断”所构成的共生系统,正是该范式未来深化研究的关键场域与核心议题。

综上,学评融合以学生核心素养的培育与发展为根本宗旨,具有三个视角的内涵。

“学”的视角。指学习者在由智能技术构建的数字化情境中,以真实任务为驱动,通过主动设计、执行与反思评价(如生成成果,成果展示、评价他人、反思改进等),并在智能体的互动与支架支持下,激发并实践高阶思维活动(涵盖知识激活与提取、迁移应用、批判性思考、创新创造、合作协作、问题解决等),从而促进核心素养的发展。它强调学习不是一个为评价做准备的过程,而是在与评价任务的深度互动中即时发生的。此视角强调学生在评价过程中的主体性、参与性与生成性。

“评”的视角。指利用智能技术对学生参与学习与评价活动的过程性表现和生成成果进行多模态、伴随式的数据采集与深度分析建模,对其在认知、社会性、心理特质等多维度的素养发展水平与状态进行智能动态分析与精准诊断,生成多维度的学习者画像与诊断报告。它强调评价的全自动、全程化和高颗粒度。此视角侧重于基于数据的价值判断与诊断反馈。

“融合”的视角。“融合”的视角是学评融合理念的核心,指“学”与“评”不再是教学过程中两个独立的环节,而是共生共存、相互嵌入、彼此促进的有机统一过程。它并非指“学”与“评”在时间或形式上的简单结合,而是指通过智能技术的介入,将两个原本独立的系统重构为一个统一的“学评共生体”。实现这种融合需要设计思维与数据智能闭环。在顶层设计上,将评价任务作为学习活动的核心载体,通过精心设计学习者与评价任务的深度交互,蕴含学习过程;在技术层面,构建一个能通过智能体实时感知、捕获学

习数据,并运用算法模型即时分析并反馈、同时又能基于此动态调整学习路径与资源推荐的智能系统。它在一个连贯的活动中,无缝整合了“作为学习的评价”(学生主体参与)、“为了学习的评价”(即时数据反馈)与“关于学习的评价”(精准素养认证),最终形成一个自我驱动、持续优化的教与学新生态。

### (三)学评融合的特征

学评融合的实践形态主要展现出以下五个维度的根本属性:人本性、智能性、功能性、融合性与积极性。下文将逐一阐释这些特征如何系统性地重塑了学习与评价的关系,并推动教育范式的深刻变革。

#### 1. 人本性(Learner-Centeredness)

学评融合强调人的核心价值和作用,评价的根本目的在于促进学习者的思维深度和核心素养的全面发展。学习者是评价思维活动的核心主体,无论是评价的学习性功能(如参与、反思)还是诊断性功能(如反馈、应用),均需充分体现学习者的主体地位。学习者是评价过程的主动设计者与参与者,主动参与设计评价、参与评价标准制定、参与评价工具研制、评价信息收集、评价数据分析处理、评价结果的撰写、评价结果的应用等全流程,评价思维和评价的主体地位充分体现。学习者评价过程是积极主动的,评价过程应设计为积极的、支持性的体验,使学习者能够获得成就感与效能感。最终评价思维过程和结果能够找到每个学习者的特点和优点,为学生的优势个性发展提供沃土。

#### 2. 智能性(Intelligence-Enabled)

学评融合的实现高度依赖于AI驱动的智能技术环境。智能技术环境不仅要支持评价的学习性与诊断性功能,更要致力于融合学习过程行为,实现智能诊断。评价的学习性功能(如成果展示、交互评价、反思引导)在基础互联网环境下即可实现学生的评价思维训练,但引入智能代理(AI Agents)、数字人等AI技术可显著提升评价思维的深度与广度,助力学生高质量、深交互、最符合最近发展区的评价过程,让自主学习与探究、交流与表达、评价过程与评价反思等环节更加智能化、个性化和常态化。评价思维过程和学习过程产生海量行为数据,以此为基础,可依托智能技术构建核心素养评价模型与学生核心素养常模,从而实现实时、智能的诊断,该诊断系统能依据不同功能需求,动态生成个性化资源、快速调节反馈信息并优化教育管理,最终为个性化的学、教、管提供精准支持。诊断性功能要求对复杂、多源、实时的过程性与结果性数据进行高效处理与分析,这必然依赖于先进的智能分析技术(如机器学习、自然语言处理、多模态数据分析)与教育心理测量技术。



### 3. 功能性(Functional Diversity)

学评融合强调评价学习性和诊断性双重功能的融合,即不分彼此,你中有我、我中有你,评价即高阶思维的学习、学习即高阶思维的评价。学习性功能是指评价活动本身即构成一种高阶学习体验,旨在激发思维、促进知识建构与能力生成,服务于核心素养的即时发展。诊断性功能是指基于过程数据对学生素养发展状态进行实时监测、精准识别与深度分析,为学习改进、教学调整及个性化支持提供科学依据。两种功能相互依存,共同作用于素养发展目标。

### 4. 融合性(Integration)

学评融合是过程与结果的协同融合、**双向**。学评融合在数字化教育生态中,同时发挥评价过程的学习性价值与评价结果的诊断性价值。评价过程不仅为高阶思维活动与核心素养发展提供载体(服务于“学”),同时也为诊断性评估的信度与效度奠定数据基础(服务于“评”)。诊断结果不仅是对学习成效的度量(服务于“评”),更通过及时、精准的反馈,为学习的持续进阶、个性化发展路径规划提供数据支撑与行动指南(服务于“学”)。二者在时空、目标与价值上实现深度交织与共生互促。

### 5. 积极性(Positivity)

学评融合的全过程都是基于学生的最近发展区,提供AI辅助预习、AI辅助生成、AI辅助评价、AI辅助反思等,目的是为了基于学生最近发展区,让**孩子**能够得到,同时又给**孩子**资源上、工具上和心理上的支持,让学生高质量的完成任务创作,获得积极的成功体验。最后也最为重要的是,学评融合是基于学生的特长发展的,所有的内容都是扬长的基础上发展。

## 三、学评融合的独特性

学评融合并非对传统评价理念的简单改良,而是在智能技术深度赋能下,对学习评价关系进行的根本性重构与范式革命。

(一)主体重构:从“被评者”到“评价思维者”——核心范式的根本转变

学评融合的核心转变在于:评价从一种外部施加的测量手段,升维为学习者内在的评价思维。这一转变的关键在于角色转换——在智能技术与智能体构建的环境中,通过系统性的活动设计,让学生参与评价标准制定、任务开发、同伴互评等环节,使其从被动接受评判转为主动进行评判。在此过程中,他们就不再是答题者,而是化身为“评价者”,这一角色转换,系统地引导并外化了一套高阶认知过程——即评价思维。它如同批判性思维与创造性思维,是学习者在审视、判断、改进自我与他人成果时所必须调

用的核心能力。它不再是教师的教学策略,而是学习者自身的一种可迁移的核心素养。因此,学习者完成了从知识接收的“客体”到运用评价思维进行自我建构“主体”的根本性跃迁。

(二)过程重构:从“线性序列”到“循环共生”——学习与评价的界限消失

在此范式中,学习与评价的二元对立关系被彻底解构,二者融合为一个连续的、内循环的共生过程。传统评价均将学习与评价视为两个可以分离的阶段或活动。然而,学评融合揭示了一个内在逻辑:学习产生评价证据,证据驱动评价思维,评价思维直接催生新的学习。这三者构成了一个不可分割的、内循环的连续体。在智能体的全程伴随下,一切学习活动过程均被记录,产生评价的证据,而证据又直接作为评价思维的输入,清晰地表明了学习过程本身就是生成评价数据的过程。这意味着,学习与评价不再是接力赛,而是同一枚硬币的两面,界限的消失使得“评中学”和“学中评”成为常态。

(三)数据重构:从“结果快照”到“全景证据”——评价数据的全面融合

学评融合通过智能体与环境驱动的多模态感知与数据融合技术,实现了评价数据维度的全面拓展与深度融合。传统的“关于学习的评价”极度依赖“结果性数据”(如考试分数),而“为了学习的评价”虽关注过程,但数据收集往往零散、主观。学评融合则借助自然语言处理、行为分析等智能算法,将过程与结果深度融合,“过程性证据”(如课堂参与度、讨论质量、草稿迭代)和“结果性证据”(作业、考试)被同等重要地置于“评价思维介入点”,共同作为思维加工的材料。这体现了过程数据和结果数据都用来做诊断的全面性。这种全息式的数据观,使得对素养的评估超越了单一的快照,实现了贯穿始终的、立体的“全景诊断”。

### (四)驱动重构:

从“外部反馈”到“内在认知驱动”——评价思维的输出直接驱动学习

学习者通过评价思维产生的三种内在认知结果(诊断性、学习性、成长性),不再是简单的报告或反馈,而是直接转化为新一轮、更高阶的学习活动。这一过程具体表现为一个连贯的认知与行动链条:学习者首先生成诊断性输出,即对自我或同伴当前学习状态(如知识漏洞、逻辑缺陷)的清晰判断,这精准定位了学习的“当前位置”;紧接着,学习性输出被激活,它特别强调了“激发元认知”,这意味着评价思维帮助学习者理解自己的学习过程,从而能够自我计划、监控、调整;而在此循环中积累的成长性输出,即对自身能力进步的觉察与自信,则提供了持续进化的内在动力。最终,这些由评价思维孕育的、关于“问题何



在“该如何做”以及“我能否做到”的深刻洞察，不再停留于纸面反馈，而是基于智能环境敏锐捕捉，动态生成或推荐更具挑战性的学习任务、提供必要的支架资源，从而驱动学习在迭代中实现质变。这是“学会如何学习”的最高体现，也是其能驱动“更高阶学习”的原因。

(五)系统重构：从“静态周期”到“进化螺旋”——强调循环与进化

学评融合构建了一个在智能算法驱动下自我强化、螺旋上升的进化系统。与传统评价的线性周期不同，其核心在于一个“数据智能闭环”：一轮评价思维所带来的洞察被智能系统即时捕获并投入应用，用于优化学习路径与任务序列，催化出在质量、深度和主动性上都更优的下一轮学习。这确保了整个系统在智能技术的调节下，学习过程不是一个简单的重复，而是一个证据质量持续提升、思维层次不断深化、素养水平持续进化的向上螺旋。

#### 四、学评融合的应用模式

学评融合理念的最终旨归，在于将“评”与“学”的共生关系付诸实践。需要明确的是，这并非将统一的理念割裂为“学习”与“考试”两个独立部分，而是同一核心范式在不同场景下的侧重性应用。二者共享着“评价即思维”“过程即证据”“反馈即驱动”的底层逻辑，共同构成了学评融合完整的实践图谱。具体而言，其应用主要归结为两大路径：其一，用在个性化学习上的学评融合；其二，用在高阶思维测评上的学评融合（如图3所示）。两者都是在智能环境赋能的基础上，实现学习与诊断两大功能的融合共生。用在个性化学习上的学评融合是以学习者评价主体，学习者在真实的评他活动中，以知识技能为载体生成、培养核心素养；用在高阶思维测评上的学评融合是以学习者评价客体，学习者在构建的问题情境中提取、运用已有核心素养解决问题，素养得以外显、量化为可测量的知识技能，

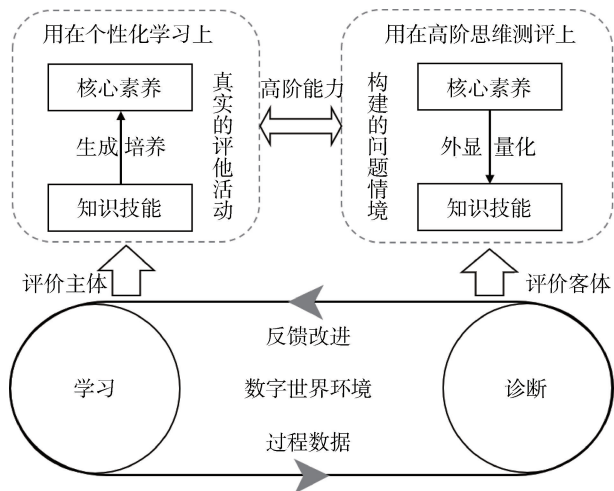


图3 学评融合的两种应用

进而实现素养的精准诊断。

#### (一)用在个性化学习上的学评融合

用在个性化学习上的学评融合理念(如图4所示)，强调评价活动的学习价值，评价活动是学习过程中的核心环节，是学习者思维的外化，其核心在于依托智能技术构建一个以学习者为主体的评价生态系统。在此场景下，生成式人工智能与智能体不再是简单的工具，而是作为核心赋能者，通过动态生成情境化任务、提供互动式学习支架、搭建协作平台并引导元认知反思，将评价活动深度内化为素养生成的过程。其核心路径表现为：学习者不再是知识的被动接受者，而是评价活动的主动设计者与参与者。他们在与智能体的深度互动中，经历“生成与分享学习成果”“交互评价同伴作品”“基于多元反馈进行反思与迭代”的完整循环。此过程倡导从知识提取到知识生成、从消极被评到评价学习的主体、从一般认知评价到高阶认知和非认知的融合。同时，伴随的过程记录和诊断是为学习服务，诊断核心除了一般的认知评价外，更强调对高阶思维能力和非认知能力的评价，如自主学习能力、合作能力、评他能力、批判性思维、创造性思维以及好奇心、坚毅、学习态度、学习动机、人格、同伴关系、身心健康等。

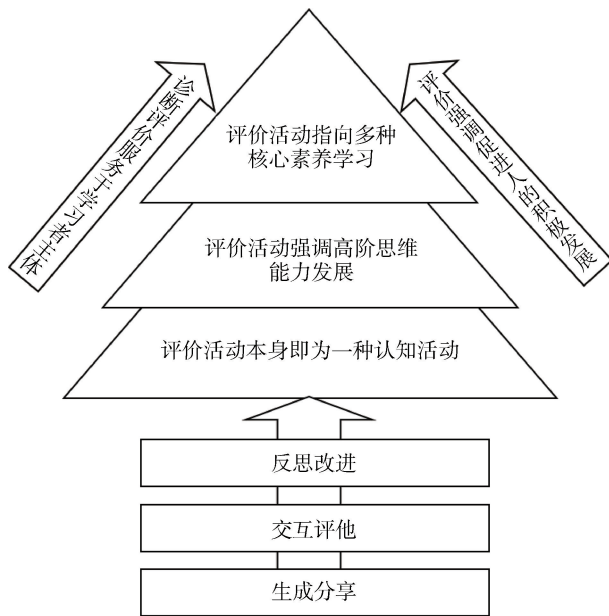


图4 用在个性化学习上的学评融合

在学评融合的个性化学习应用中，智能技术系统性地构建了一个“生成—评他—反思”的闭环学习流程，使学生在深度参与中实现素养的持续生成与发展。首先，在“生成与分享”环节，学习者在智能创作工具与生成式AI的协作支持下，完成从研究报告到创意作品等多种形态的成果输出。AI不仅提供资源与灵感，更通过多轮对话挑战学生思维定式，深化其对知识的理解与创造性运用。其次，在

“交互评他”环节,智能体通过评价支架与自然语言处理技术,引导学生关注逻辑严谨性、证据充分性等关键维度,并对其评语质量进行分析与提示,将简单评判升华为高质量的评他思维训练。最终,在“反思改进”环节,基于全过程的多模态数据分析,系统生成个性化的“素养发展画像”,并以可操作的方式呈现学生的优势、短板与成长轨迹,在智能引导下直接转化为改进建议与学习资源,驱动学习策略与作品的持续优化,实现素养的螺旋式进阶。

评价的学习性是从评价的学习意义出发,评价过程即是学习的发生过程,评价结果促进学习维持与个体发展。首先,评价活动本身即为一种认知活动,学习者在进行自评或他评的过程中,主动分析判断自我和他人作品的优缺点,这一过程就是认知与元认知能力的实际应用,可以在学习活动的任何环节或活动中。其次,评价活动强调高阶思维能力发展,评价活动不仅是一般的认知、元认知的学习活动,运用批判性思维去审视他人作品的逻辑性、创新性与有效性,这需要评他能力、评价反思能力等高阶的认知能力,属于高阶的认知活动。再次,评价活动指向多种核心素养学习,评价的全思维过程不仅是促进人评价思维发展的载体,也通过小组合作评价任务,促进学习者沟通协作、情绪调节与人际交往等核心素养发展。四是评价的学习性强调对人的发展,是进阶进步的、积极的、正向的、是自信的,多元的,是有成功体验的。五是学习过程的诊断评价服务于学习者主体的发展需求,而非将学习者视作被动评价的客体,通过数字平台伴随记录的学习与评价过程数据,建立时序过程和总结性分析模型,诊断学习者认知、社会性、心理等发展状况,再通过可理解的方式可视化,反馈学习者与管理者等,促进学习过程的调节与改进。

## (二)用在高阶思维测评上的学评融合

用在高阶思维测评上的学评融合理念(如图5所示),在以智能体与环境为核心技术支撑的测评系统中,强调对学生核心能力的诊断价值,其核心在于借助智能技术实现评价诊断的精确性与科学性。核心过程一般包括背景介绍、资料学习、探究实验、问题解决、反思等环节,智能体与环境在其中扮演着情境构建、过程引导与数据感知的关键角色,每个环节伴随的过程记录和问题解决的结果是评分的重要依据。

在测评中,学评融合采用“先学后诊”的模式,聚焦学习者高阶认知能力和非认知能力的评价,通过生成式人工智能技术,可动态构建高度仿真的复杂问题情境,例如模拟科学探究、工程设计或社会决策等场景,以考察学习者的科学探究能力、自主学习能力、合作能力、批判性思维、创造性思维、数字世界学习能力、AI素养等。测评工具开

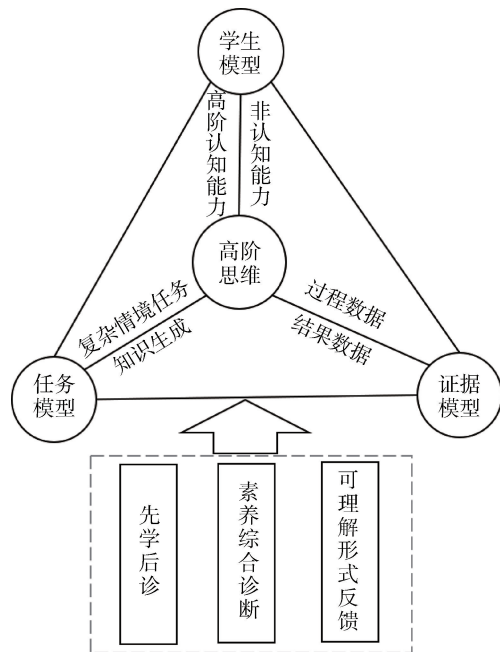


图5 用在高阶思维测评上的学评融合

发通常采用证据中心设计模式(ECD),通过明确设计学生模型、任务模型与证据模型三个核心模块,以支持设计复杂任务并收集和分析过程数据,进而实现对交互环境中复杂高阶能力的概念化与评估。具体的操作体现在以下几个方面:一是借助智能任务生成引擎,强调运用高阶思维进行问题解决的过程评估。测评不再使用固定题目,而是由系统动态生成数字情境任务,学习者运用已有的核心素养,快速识别问题、探究规律、形成新知识和解决方案,测评着重考察学生如何使用高阶能力解决问题,而非对原有记忆知识的提取。二是通过智能模拟环境设计复杂情境任务以考核能力应用。测评中学习者需要参与复杂的科学探究或工程设计任务,通过自主学习、科学实验、合作探讨生成新知识或解决方案。在此过程中,智能体作为交互对象与环境要素,实时采集学生的操作序列、决策逻辑、协作对话等多维行为数据。三是基于智能算法模型实现多源数据的融合分析与能力诊断。系统通过计算机视觉、自然语言处理等技术,自动采集学习者在开放、交互的模拟环境中的海量行为数据,并利用机器学习方法建立从行为证据到能力维度的映射模型,以支持对参测者能力的统计推断。四是通过学生模型指导任务开发、依托证据模型整合过程性数据与结果性数据。最后,该模式的核心逻辑在于构建一个智能驱动的“能力诊断—反馈”闭环。基于人工智能的评分引擎将表现学习者能力的过程性行为数据(如操作路径、策略选择)与生成性成果(如作品、解决方案)作为评分依据,并据此制订证据基础,评估结果从答案导向转向证据驱动,从而实现对动态高阶能力和学习增值的全面评估。



此处的学习过程强调对学习者的\*\*高阶思维评价\*\*、学习者是\*\*被评价的客体\*\*。

### 五、学评融合的实践策略

随着生成式人工智能等数字技术的快速发展,发展新质生产力亟需大批在智能环境中进行创造、协作与解决问题的复合型人才。如何快速推进教育评价变革,使学评融合理念落地,这不仅关乎理论认知,更取决于能否转化为行之有效的实践方案。为跨越从理念到行动的鸿沟,系统性、可操作的策略指引至关重要。基于前述学评融合的概念、内涵、特征、独特性与应用模式,下文将从目标、文化、系统与人才四个层面,提出一套协同推进的实践策略,以期教育工作者与决策者提供清晰的行动框架,共同推动评价生态的根本性变革。

#### (一)坚持素养导向,推行渐进式教育评价改革路径

学评融合的推进应采取渐进式改良路径,在现有体系基础上实现由智能技术驱动的评价功能的“赋能式”升级。具体而言,应在保留分数等传统指标合理功能的同时,借助智能算法对学习过程进行多维度分析,生成融合知识掌握、思维过程、协作表现等的“素养画像分数”。OECD在《教育2030:学习指南》中强调利用分数等传统工具的“转型再设计”,体现创造力、批判性思维、解决问题等素养<sup>[8]</sup>。通过“赋能分数”引导公众从追求成绩向重视素养转变,形成改革的社会共识。同时,推进评价工具从结果筛选向过程支持的内核转型,要积极开发基于生成式AI与智能体交互的表现性评价任务,使素养评价在课堂层面落地。融合国际经验与数字技术,发展面向未来素养的个性化评价工具。通过双向改革路径,逐步构建服务学生发展、体现中国路径的教育评价体系。

#### (二)重塑评价文化,构建以学习者为中心的评价生态

学评融合回应了当下教育评价实践中普遍存在的结构性问题,是面向行动的实践型理论创新。当前评价实践受管理主义思维主导<sup>[9]</sup>,导致“唯分数”“唯升学”等倾向普遍存在<sup>[10]</sup>。必须从根本上扭转管理主义导向的评价文化,将评价重心从外部控制回归至育人本质,构建以学习者为中心的协同评价生态。构建新型评价生态,需推动评价主体从单一走向多元,教育者应系统地将学生自评与同伴互评嵌入学习全过程,引导他们在参与标准制定、作品评价与反馈交流中发展评价思维与元认知能力。同时,需充分利用智能技术对学习过程进行伴随式记录,生成可视化学情报告,为教师调整教学与学生自我优化提供精准依据。学校管理者亦需建立支持性评价观,将评价结果用于教学改进与资源优化,而非简单用于绩效评判,从而营造具有安全感、信任感与成长性的评价氛围。

(三)推动全员参与,构建多主体协同联动的评价体系  
学评融合的成功实施有赖于学生、教师与管理者三类主体的协同发力与系统联动。学生能力结构在新质生产力推动下持续重构,也对教育支持系统提出更高要求。教师作为关键连接者,需从知识传授者转向“人机协同”模式下的学习设计师与评价引导者。这一角色转型要求教师具备掌握设计与解读AI赋能的评价任务、运用智能工具开展差异化指导的能力。与此同时,教师还需主动营造互动评价氛围,激发学生的参与感与主体性。在新时代教育评价改革中,公共理性教育的重点在于助力评价主体和治理主体树立权力、责任和利益相统一的教育评价理念,以明确区分自身的价值诉求与其负载的其他主体的价值诉求<sup>[11]</sup>。因此,教育管理者作为制度塑造者,其评价观也需从控制导向走向支持导向。教育管理者应利用智能治理平台,建立数据驱动的发展性教师评价机制,从多个维度评估教师在促进学生素养发展方面的贡献。通过构建贯通“学生—教师—管理者”的智能评价联动系统,实现基于统一数据底座的多主体协同与育人生态的整体优化。

(四)加强专业支撑,系统培养评价领域的复合型人才  
为支撑学评融合的规模化、专业化发展,必须加快建设懂教育、通评价、善用技术的复合型人才队伍。建议设立教育评价相关二级学科与专业,开设融合智能教育技术、教育数据挖掘、人机协同评价设计、教育理论、测量统计的跨学科课程,培养具备评价设计、数据分析与政策理解能力的专业人才。针对在职教师、教研员与教育管理者,应开展分层分类的专项培训,重点提升其“人机协同”能力,包括驾驭智能评价工具、解读AI生成的分析报告、以及在人机协作环境中引导学生发展的实践智慧。同时,应积极建设跨校、跨区域的数字化教研平台与专业学习社群,通过AI辅助的案例分析与线上协同研讨,形成常态化、网络化的专业支持体系,为学评融合的持续深化提供坚实的人力基础与智力保障。

### 六、结语

面向教育强国建设的新征程,深化教育评价改革已不仅是教育系统的内在要求,更是关乎国家人才战略与未来竞争力的关键举措。传统的评价体系因其难以胜任核心素养的评估、难以摆脱“唯分数”的惯性而面临严峻挑战<sup>[12][13]</sup>,构建一个既能精准甄别又能有效育人的新型评价范式,已成为当前教育改革的紧迫任务与核心突破口。本文系统阐述的“学评融合”理念,正是在生成式人工智能、智能体等智能技术构成的新型数字底座支撑下,为回应这一时代命题所提出的理论框架与实践路径。

展望未来,学评融合的研究与实践仍处于探索阶段,



尚有广阔的发展空间。未来的研究可重点关注以下几个方向：其一，理论深化，进一步构建和完善具有中国特色的学评融合理论体系，并探索其与不同学科、学段深度融合的特殊规律；其二，技术赋能，亟需研发更先进的生成式AI教育应用、智能体互动机制与多模态数据分析技术，攻克高阶思维与非认知素养测量的难题；其三，教师发展，需要探索人机协同的教学与评价新模式，培养教师在智能环境下的专业能力。

#### 参考文献：

- [1] 习近平主持召开新时代推动东北全面振兴座谈会强调牢牢把握东北的重要使命奋力谱写东北全面振兴新篇章[N].人民日报,2023-09-10(01).
- [2] 中共中央 国务院印发《教育强国建设规划纲要(2024-2035年)》[EB/OL].  
[http://www.gov.cn/zhengce/2020-10/13/content\\_5551032.htm](http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm),2025-01-19.
- [3] 中共中央 国务院印发《深化新时代教育评价改革总体方案》[EB/OL].  
[http://www.gov.cn/zhengce/2020-10/13/content\\_5551032.htm](http://www.gov.cn/zhengce/2020-10/13/content_5551032.htm),2020-10-13.
- [4] 檀慧玲,王玥.论教育数字化进程中教育评价理论重塑与模式创新[J].中国教育学刊,2025,(3):55-61.
- [5] 骆方,田雪涛等.教育评价新趋向:智能化测评研究综述[J].现代远程教育研究,2021,33(5):42-52.
- [6] 张生,王雪等.人工智能赋能教育评价:“学评融合”新理念及核心要素[J].中国远程教育,2021,(2):1-16.
- [7] 张生,李文硕等.基于“学评融合”理念的个性化评价支持环境设计及其应用[J].中国电化教育,2023,(7):91-98+117.
- [8] OECD.OECD Learning Compass 2030[M].Paris:OECD Publishing,2019.
- [9] 朱德全,冯丹.新时代教育评价改革多元主体的行动逻辑[J].现代教育技术,2024,34(5):5-15.
- [10] 易凌云.“五唯”问题:实质与出路[J].教育研究,2021,(1):4-14.
- [11] 曹渡帆,朱德全.新时代教育评价改革的数字正义何以可能[J].电化教育研究,2023,(12):21-27+34.
- [12] Stiggins,R.J.From formative assessment to assessment FOR learning:A path to success in standards-based schools[J].Phi Delta Kappan,2005,87(4):324-328.
- [13] Au,W.High-Stakes Testing and Curricular Control:A Qualitative Metasynthesis[J].Educational Researcher,2007,36(5):258-267.

#### 作者简介：

张生：教授，博士，研究方向为教育评价与智慧测评。  
王雪：在读博士，研究方向为教育测量与评价。  
辛涛：教授，博士，研究方向为心理测量与教育评价。  
梁凯丽：二级教师，博士，研究方向为教育测量与评价。

## The Concept, Connotation, and Application Models of Learning-Assessment Integration in the Age of AI

Zhang Sheng<sup>1</sup>, Wang Xue<sup>1</sup>, Xin Tao<sup>1</sup>, Liang Kaili<sup>2</sup>

- 1.Collaborative Innovation Center of Assessment for Basic Education Quality, Beijing Normal University, Beijing 100875  
2.Beijing Xicheng District Academy of Education Sciences, Beijing 100035

**Abstract:** The AI era has brought forth new demands for educational assessment. Traditional evaluation systems, which focus primarily on knowledge and skills while emphasizing screening and selection, are increasingly inadequate for meeting the strategic goal of developing students' core competencies. This paper proposes and systematically elaborates on a novel concept and practical framework termed the "Integration of Learning and Assessment." The study begins by defining the core connotation of this integration: it transforms assessment tasks from passive exercises for learners into evaluative learning activities that are actively participated in by learners and empowered by AI agents. Leveraging smart technologies, it forms a closed loop in which "assessment embeds learning, learning data informs assessment, and assessment outcomes drive iterative learning." Furthermore, the paper analyzes its internal structure and characteristics from three lenses: "learning," "assessment," and "integration." The research further reveals five fundamental attributes that distinguish the Integration of Learning and Assessment from traditional evaluation concepts. On this basis, the paper clarifies its dual-track application model in personalized learning and the assessment of higher-order thinking, and ultimately proposes four practical strategies centered on competency orientation, cultural reshaping, system-wide collaboration, and professional empowerment. The concept of Integration of Learning and Assessment not only provides a theoretical anchor for addressing the current challenges in educational assessment reform but also, through its systematic implementation pathway, points a forward-looking direction for constructing a new assessment ecology that empowers the holistic development of students and aligns with the context of Chinese education.

**Keywords:** learning assessment; era of AI; the integration of learning and assessment; evaluation reform

收稿日期：2025年5月30日  
责任编辑：邢西深