

# 新时代教育增值评价实施现状、困境及改进路径

方嘉玲

杭州师范大学经亨颐教育学院, 浙江 杭州

收稿日期: 2026年6月3日; 录用日期: 2026年7月3日; 发布日期: 2026年7月8日

## 摘要

本文系统考察了《深化新时代教育评价改革总体方案》实施以来我国教育增值评价的实践进展, 深入剖析了当前面临的现实困境, 并提出了系统化的改进路径。研究发现, 增值评价已在宏观区域、中观学校及微观课堂层面形成了多层次、多维度的实践探索, 其理念正从学业成绩增值向综合素质的全面发展拓展。然而, 实践过程中仍面临多重挑战, 包括育人初衷存在被功利主义异化的风险、复杂统计模型与教育现实之间的适配难题、数据采集分散和结果应用存在“最后一公里”困境、系统协同不足与“唯分数论”的文化惯性。为此, 本文提出应从强化顶层设计与协同机制、构建“五育并举”数据体系、建立诊断性反馈与支持闭环、推动评价文化根本转型四个方面进行系统性改进, 以构建回归育人本源、技术科学可行、实践落地有力的增值评价新生态, 为深化教育评价改革、推动教育高质量发展提供参考。

## 关键词

教育增值评价, 实施现状, 现实困境, 改进路径, 教育评价改革

## Current Implementation Status, Challenges, and Pathways for Improvement in Value-Added Assessment in Education in the New Era

Jialing Fang

Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou Zhejiang

Received: June 3, 2026; accepted: July 3, 2026; published: July 8, 2026

## Abstract

This paper systematically examines the practical progress of educational value-added assessment in China since the implementation of the “Overall Plan for Deepening Educational Evaluation Reform in the New Era”, provides an in-depth analysis of the current practical challenges, and proposes a systematic path for improvement. The study finds that value-added assessment has given rise to multi-level, multi-dimensional practical explorations at the macro-regional, meso-school, and micro-classroom levels, and that its philosophy is expanding from academic achievement gains to the comprehensive development of overall quality. However, the implementation process still faces multiple challenges, including the risk of the original educational intent being distorted by utilitarianism, difficulties in adapting complex statistical models to educational realities, the “last-mile” dilemma in data collection and application, as well as insufficient systemic coordination and the cultural inertia of “score-centric” evaluation. To address these issues, this paper proposes systematic improvements in four key areas: strengthening top-level design and coordination mechanisms; constructing a data system that supports the “five-fold education”; establishing a closed-loop system of diagnostic feedback and support; and promoting a fundamental transformation of the evaluation culture. These efforts aim to build a new ecosystem for value-added assessment that returns to the fundamental purpose of education, is technically and scientifically feasible, and is effectively implemented in practice, thereby providing a reference for deepening educational evaluation reform and promoting high-quality educational development.

## Keywords

Value-Added Assessment in Education, Current Implementation Status, Practical Challenges, Pathways for Improvement, Reform of Educational Assessment

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

教育评价直接关系到教育的发展方向，更决定着教育强国建设的成败。长期以来，我国传统的教育评价模式侧重于结果性与终结性评价，而这种以分数和升学率为核心的绩效主义导向，不仅容易忽视学生、教师与学校的初始基础与背景条件，加剧教育焦虑与无序竞争[1][2]，更有悖于教育公平与立德树人的根本任务。为了扭转这种不科学的评价导向，破除“五唯”顽疾，国家层面开始大力倡导增值评价。2020年，中共中央 国务院印发《深化新时代教育评价改革总体方案》<sup>1</sup>直接从顶层设计上明确提出“探索增值评价”，将其作为改进结果评价、强化过程评价、健全综合评价的“四维”改革布局中的关键一环[3]。随后，2022年版的《义务教育课程方案》再度强调，要积极探索增值评价，重点关注学生在过程中的真实进步[4]。

这一系列政策出台以来，我国学者对传统评价模式进行深刻反思，并结合本土实际开展了一系列增值评价实践探索，为新时代评价改革提供了丰富案例和多元视角。本文基于对方案实行后增值评价实施现状的系统考察，聚焦评价推行的现存困境和改进路径，旨在厘清三个关键问题：一是当前增值评价实

<sup>1</sup>[https://www.gov.cn/zhengce/202203/content\\_3635463.htm](https://www.gov.cn/zhengce/202203/content_3635463.htm)

施现状如何？二是增值评价实施以及进一步推行过程中，还面临哪些现实困境？三是基于上述讨论，未来应如何系统性构建增值评价发展的改进路径？以期为深化教育评价改革、推动教育高质量发展提供参考性建议。

## 2. 增值评价实施现状

自2020年《深化新时代教育评价改革总体方案》将“增值评价”提升至国家战略高度以来，我国教育界对增值评价的探索迈入了以实践应用为核心的全新阶段。截至2025年12月，各地实践以前所未有的广度和深度展开，逐步形成了一幅多层次、多维度、多主体，从宏观区域到微观课堂的实践图景。

### （一）宏观区域层面的协同推进

在区域教育实践中，增值评价已成为推动教育生态优化、促进公平与质量提升的重要工具。各地基于自身的数据条件和改革目标，发展出了各具特色的实施路径[2]。这些实践的一个共同趋势，是超越了传统的学校排名，转而关注教育生态的整体改善。例如，湖南长沙长期坚持“从入口看出口、从基础起点看发展变化”的评价思路[5]，四川成都武侯区则构建了以“进步值”为核心的“双驱三核”模型，旨在将教育工作者的视线引向学生和学校的动态进步，并引导全区树立关注成长的教育评价观[6]。更值得关注的是，增值的理念正在向学生综合评价领域延伸，一些地区正积极探索如何衡量学生更广泛的成长。山东济南的区域实践涉及对学生学业和身心健康水平的增值分析[7]，浙江温州推行“初中绿色增值评价”，将品德发展、学习生活幸福感等纳入考查范围，试图对“五育”的进步进行量化考查[8]，这些探索表明，增值评价正从单一的学业成绩增值，转向更加综合的全面发展评估。随着技术进步，一些地区还开始借助大数据和可视化技术，将复杂的增值数据转化为清晰的诊断报告，让评价结果更加直观，从而能够更精准地赋能教学改进[9]。

### （二）中观学校层面的校本化创新设计

在学校与课堂的具体情境中，增值评价的实践重点则聚焦于具体的教育教学改进。四川成都棕北中学西区实验学校的探索极具代表性。该校没有直接套用通用模型，而是自主研发了包含学业、品德、行为、健康等多维度的校本指标体系，并融合多种评价模型。这套量身定制的体系，其价值不仅在于评估，更在于它能预测学生的“合理发展区间”，精准定位个体在不同学科上的潜力，最终转化为优化教学与资源的科学参照[10]。如何让“增值”这一概念变得可知可感？山东省菏泽市定陶区第二实验小学的实践提供了另一种思路。他们以“让学生看见自己的成长”为核心，创设了“四横四纵”全景式评价。这种多维度、过程性的记录方式，旨在清晰描绘学生的成长轨迹，从而激发其内在动力[11]。这些案例共同表明，校本化、富有创意的评价设计，能够更好地服务于个性化的学生发展和精准化的教学诊断。

### （三）微观课堂层面的学科教学深度融合

增值评价在学科教学，尤其是语文写作、科学、信息科技等强调过程与能力发展的领域，展现出巨大的应用价值。在语文写作教学中，有研究者通过设置不同课型，将增值理念贯穿于“自能修改”“互评互改”等环节，使评价成为驱动学生写作能力螺旋上升的内在动力[12]。还有学者将增值评价融入小学科学的学习中，提出借力增值的评价策略，并以学生的需求和发展为主导，建构了“实景动评-结构智评-项目创评”的增值评价体系，将学生隐形的科学素养的提升直观化和显性化。例如在《身体的运动》课中复现真实情境，通过模型操作与表达对学生进行动态评价[13]。在信息科技等学科中，增值评价也被用于追踪学生核心素养的达成过程，通过设置真实的评价任务，诊断和激励学生持续进步[14]。通过以上案例可以发现，增值评价在微观课堂层面的应用使得教学评价从原本静态结果的评判转向动态过程的关注，让学习过程中难以量化的素养提升变得可视、可感，进而使评价本身成为推动学生持续发展的重要力量，真正体现“以评促学、以评促长”的教育理念。

### 3. 增值评价的现实困境

1990年代, Sanders和Horn开发了田纳西增值评价系统(TVAAS), 基于田纳西州完善的年度标准化测试数据, 开创性地将混合模型方法论应用于教师效能评估[15][16]。

然而, 经过多年的发展和演变, 增值评价仍面临持续的挑战与争议。2010年, Rothstein通过精巧的“伪造检验”揭示标准增值模型存在学生非随机分配导致的偏误, 动摇了增值评价无偏估计的核心假设[17]。Darling-Hammond等学者进一步指出教师增值排名的年际波动显著, 不适合作为高利害决策的依据[18]。这些批评最终汇聚为两份重要声明: ASA指出教师仅解释学生成绩变异的1%~14%, 警告不应将增值评价作为高利害决策的单一工具[19]; AERA则提出了采用增值模型前应满足的八项技术标准[20]。

在国际比较视野下, 中国推进增值评价同样面临若干独特的结构性困境, 这些困境贯穿于价值理念、技术模型、数据采集、结果应用乃至整个教育生态系统之中。

#### (一) 育人初衷与实践异化矛盾

增值评价的初衷在于回归育人本质, 但在实践过程中常被模糊理解甚至错误解读, 存在理念认知不足和实践异化的风险。

其一, 认知不足, 导致理念与实践脱节。尽管国家政策文件中多次倡导探索增值评价, 但多为宏观的原则性表述, 缺乏清晰的操作定义与系统的实施指南[1]。这使得增值评价仍主要停留在学术讨论层面, 尚未转化为一线教育工作者所能理解和运用的实践模式。政策理念层面的倡导与教育现场的实际操作之间存在显著脱节, 许多教师和管理者对增值评价的概念内涵、实施目的及具体方法缺乏清晰认知, 甚至存在理解偏差。正因如此, 增值评价面临边缘化、表面化的风险, 难以真正深入教育教学的日常实践中。

其二, 功利主义取向仍然盛行, 导致评价目标窄化与异化。在强大的应试文化惯性中, 增值评价以“发展”为核心的理念, 往往在实践中被异化, 再度走向以“甄别筛选”为目的的传统路径。实践中, 许多探索不自觉地将“增值”简单等同于“分数增长”, 评价内容高度集中于可量化的学业成绩, 而学生的品德、身心健康、实践创新能力等核心素养则因“难以测量”而被边缘化[2][21]。更严重的是, 一旦将增值评价结果与绩效、排名、资源等利害挂钩, 其本应具备的诊断、改进功能便被严重削弱。学校和教师的行为逻辑可能从“如何促进学生全面发展”异化为“如何获取更高的增值分数”, 使增值评价沦为又一种竞争与排名的工具[2]。这种异化使增值评价背离了育人初心, 有重蹈“唯分数”覆辙的风险。

#### (二) 理想模型与复杂现实的背离困境

在国际学术界, 围绕增值评价模型(Value-Added Models, VAMs)的争议源源不断。支持者认为增值评价能够有效剥离学生背景因素, 更公平地反映教师与学校的教育贡献。2014年, Chetty等学者的大规模追踪研究即表明, 优质教师的增值效果对学生长期收入具有显著正向影响[22]。然而, 批评者也对此提出了一定的质疑。ASA指出增值分数受诸多不可控因素影响、标准误偏大; AERA则警告将增值结果用于高风险教师评价决策存在方法论风险。

不可否认, 技术模型是增值评价的科学支柱, 但从国际视野和国内实践来看, 它仍然存在一定的自身局限, 难以完全适配教育的复杂性, 这就构成了我国增值评价实施中的又一核心困境。

首先, 模型选择与应用黑箱难题。当前增值评价模型种类繁多, 如增益分数模型、协变量校正模型、多层线性模型等, 各有其统计学假设与适用条件[23][24]。然而, 这些模型往往专业性强、计算复杂, 对于绝大多数一线教育工作者而言犹如“黑箱”。他们既难以理解模型背后的原理与假设, 也无力在众多模型中进行科学选择和参数调试[25]。这种技术壁垒导致模型的应用严重依赖少数专家, 评价过程同教育主体相脱离, 最终严重削弱了结果的公信力与可接受度。

其次，“净效应”剥离的理想与现实困境。增值评价的核心追求是剥离学生自身、家庭背景等前期因素的影响，测量学校或教师的“净效应”。但这在本质上是一个统计学上的理想。现实中，影响学生成长的变量是非常多的，统计模型只能控制部分可观测、可量化的变量，但对于内化于学生身心的隐性文化资本、非认知能力、社会情感因素等，则难以捕捉和量化[26]。这意味着，增值评价无法获得真正意义上的“净效应”，并且在现实层面上其稳定性和一致性仍然面临严峻挑战。

最后需要正视的，是数据与测量本身固有的局限性。增值评价依赖于连续、完整的数据轨迹，然而实践中，学生流动、缺考或记录疏漏导致的数据缺失在所难免。此外，无论是标准化测试还是其他测量工具，其结果都难以完全精确地反映学生的真实水平，存在一定的测量误差[25]。这些测量误差在复杂的统计模型中会被传递和放大，进一步影响评价结果的精确性与可靠性。

### (三) 数据采集和结果应用挑战

要让增值评价体系真正运转起来，关键在于获得高质量的系统数据并加以有效利用。然而，在实际情况中，从数据采集到结果应用的过程仍然存在一定的阻碍。

一方面，数据采集面临分散、混乱、缺失等问题。学生的相关信息往往分散在不同管理部门、不同学段乃至不同信息系统中，标准不一、平台各异、共享困难，难以形成连续完整的个人发展画像[2]。另外，学业成绩数据相对集中，但更能反映学生全面发展的品德、身心、艺术、劳动等维度数据，则因涉及面广、量化难度高、主观成分多，其收集的信度和效度始终存有争议[10]。例如，综合素质评价多依赖主观填报，难以量化，其真实性也常受质疑[27]。这直接影响了评价本身的科学性和公信力。

另一方面，数据分析和结果应用也存在一定的瓶颈。运用高级统计模型处理数据后得到的增值评价报告往往充斥着统计术语和复杂结果，对于学校管理者和教师而言可读性差、理解门槛高[21]。即使理解了关于某学科增值效果低于预期的结果，也常常困惑于“为什么会这样”以及“如何改进”的问题。也就是说，针对结果的专业解读与改进建议仍然比较缺乏，即使有也是停留在描述和判断层面，未能提供清晰的行动路线图。同时，针对学生个体的增值分析极为欠缺，评价如何从对学校、教师等群体的效能评估，精准转化为对每一个学生个性化学习的支持，仍是悬而未决的难题[26]。

### (四) 协同不足与文化惯性的系统生态困境

增值评价的成功实施非单一技术问题，更是一个涉及多方主体、需要文化土壤的系统工程。从组织协同看，增值评价需要教育行政部门、教科研机构、学校、教师乃至家长形成合力。但目前，行政推动、教研支撑、学校落实之间常常脱节。研究者与实践者之间容易形成信息不对称，先进的评价理念难以转化为一线教师用得上的工具和方法[1]。从区域均衡看，增值评价对数据基础设施、专业人才、经费投入要求较高，这就导致发达地区、优质学校可能先行先试，而广大农村地区、薄弱学校则面临较大的技术门槛和实施阻碍，加剧了教育不公的现象[26]。从文化重塑看，长期形成的“以分数论英雄”的优绩文化和“以结果评绩效”的管理模式，是增值评价面临的最顽固的深层阻力。若不从根本上推动学校评价文化和治理模式的转型，任何技术层面的改进都可能被旧有文化所扭曲。

## 4. 增值评价改进路径

针对增值评价在推行过程中面临的多重困境，未来的突破不应止步于单点修补，而需推行系统化的改进路径，着力构建一种真正回归育人本源、技术支撑扎实、实践落地有力的增值评价新生态。

### (一) 强化顶层设计与多方协同，系统推进评价改革

当前增值评价实践中的一些问题，如协同不畅、政策指引不够明确等，在一定程度上导致了执行过程中的偏差与理念认知的混淆。为此，我们亟需推动治理体系的优化，构建权责分明、运行顺畅、多方参与的协同机制，实现从“政策倡导”到“系统赋能”的转变。

其一，强化国家层面的宏观指引与标准制定。教育部及相关部门应牵头制定有关增值评价详细具体的实施方案和实践指南，对其核心概念、适用范围、基本原则、模型选用规范以及结果解释与伦理红线进行清晰界定。这不仅能为各个地区的实践提供明确的行动导航，更能有效遏制将增值评价窄化为“分数增长排名”的异化倾向。同时，需构建统一的国家级教育数据平台，整合学籍管理、体质监测与综合素质评价等关键系统，以此为基础保障，赋能数据驱动的教育评价改革。

其二，构建行政、教研、学校、社会“四位一体”的协同机制。增值评价实践的专业性要求我们必须打破部门壁垒。教育行政部门负责宏观规划与资源统筹；教科研机构应转型为区域增值评价改革的智慧库，为学校提供模型选择、数据分析和解读等专业支持，切实贯通理论与实践；学校作为实施主体，应获得充分的自主权，积极开展符合校情和学情的校本化评价创新；同时，应积极引导家长与社会舆论，通过官方解读、家长课堂等形式，向社会公众普及增值性评价的理念与价值，建立对增值评价的科学认知，营造科学评价的社会氛围。

### **(二) 构建“五育并举”数据体系，分阶段、分主体推进评价改革**

大数据等现代技术的发展可为增值评价的有力实施进一步赋能，为其提供更有效、更有针对性的技术支持[28]。因此未来应战略性地利用大数据与人工智能来破解当前存在的模型“黑箱”与数据采集难题，从而在增值评价实施过程中实现数据可得、模型可释、工具可用的理想状态。

因此，构建“五育并举”的立体化数据采集体系至关重要。未来的数据收集必须超越对学业成绩的单一依赖，而应整合学业水平测试、体质健康监测、心理测量表、课堂行为观察、项目式学习作品、社会实践记录等多来源数据。特别是在品德发展、美育素养等难以量化的领域，应大力探索“量化与质性结合”的策略[29]。例如，开发基于真实任务的表现性评价工具，或利用人工智能技术对学生的课堂发言、作品进行深度分析，实现对高阶思维和核心素养发展的过程性追踪与记录。然而，这一数据体系的构建并非一蹴而就，必须正视实施过程中的现实障碍，采取分阶段、分主体的渐进式推进策略。具体推进策略如下：

第一阶段：校本数据整合试点。以部分试点学校为单位，聚焦学业成绩与行为表现数据(如课堂参与度、作业完成情况)的采集与整合，建立校级简易数据平台。此阶段主要面临资金不足例如硬件采购、软件定制成本较高和教师技术素养欠缺的挑战。应对策略包括：申请地方教育专项经费或引入低成本的开源数据工具；开展面向教师的“数据素养”校本培训，降低技术门槛。

第二阶段：区域数据互通。在试点基础上，由区域教育主管部门牵头，统一数据标准与接口协议，实现学校间数据的互联共享，并逐步纳入体质健康、心理测评等更多维度数据。此阶段技术挑战在于学生数据隐私保护以及不同学校间数据共享意愿。应对策略有：采用隐私计算与分级授权技术；制定区域数据伦理规范。

第三阶段：智能化分析与应用。借助人工智能实现可解释性增值评价，为教师提供直观的诊断画像。此阶段挑战集中于模型的可解释性与教师接受度。应对策略为：引入第三方专业机构参与模型验证与工具开发，建立教师与数据科学家的协同工作坊，通过案例示范让教师亲身体验“诊断画像”对教学改进的实际价值。

通过上述分阶段分主体的实施策略的设计，可以有效降低增值评价改革风险与实施阻力，使增值评价从“自上而下的要求”转变为“自下而上的需求”。

### **(三) 强化结果应用，破解增值评价“最后一公里”难题**

当前增值评价的结果应用仍然面临巨大的“最后一公里”问题，这严重削弱了增值评价的诊断与改进价值。因此，必须建立数据驱动到教学改进的完整闭环，形成从评估、反馈到支持的系统路径，让评价真正推动教学实践的精准改进。

一方面,要构建一个服务于多主体的诊断性反馈体系。增值评价报告不能仅仅呈现一个结果,而应该诊断出真实存在的问题,并说明未来应如何针对性改进。整个反馈体系应以学生为中心展开,首先要为学生个体提供清晰的成长反馈,让其看见自己的进步与可能,激活内在动力。而在面向教师时,应提供班级、学科及知识点的具体增值分析,使教师能将评价结果真正用于教学方案的改进。同时,还需为教育管理者呈现区域和学校层面的深度诊断,聚焦优势与短板,为区域评价改革提供指导性意见。

另一方面,需将评价结果深度融入校本教研与教师发展。学校层面可以开展一些教学反思会,组织教师共同解读增值报告,围绕报告中的重点问题进行集体备课,还可以针对性地开展课例研究。同时,教育研究部门需依据评价结果准确识别教师群体的能力短板,提供定制化的专题培训,并持续跟进指导,真正提升教师的教学改进能力。

#### (四) 以文化转型为根本,推动评价从证明走向发展

最后,也是最为根本的,是推动整个教育生态的文化转型。根深蒂固的“唯分数”绩效主义文化是增值评价实践异化的深层原因。若文化环境不变,任何技术与制度的改革都难以逃脱“新瓶装旧酒”的困境。

第一,解绑评价结果与利害奖惩,让其回归专业发展的本质。在实施过程中,必须明确增值评价的核心功能是“诊断改进”而非“筛选排名”,应尽量避免将其直接与教师绩效、学校排名、资源分配挂钩,以消解学校和教师的抵触与焦虑,从而能够更加无所顾虑地投身于评价改革实践中去。等到评价体系成熟、文化认同建立后,可审慎地将其作为综合评价的参考性依据之一。

第二,建设滋养评价改革的学校文化。学校是变革评价生态的关键场所,这要求管理者和教师有意识地在会议、校园宣传与家校沟通中,持续倡导一种重视过程、鼓励尝试、尊重差异的评价文化。另外,还可以设立“学生进步奖”、举办“成长故事会”等活动,让抽象理念变得可触摸、可参与,让师生在日常中感受到进步被看见、过程被珍视,从而使增值理念成为师生认同并践行的日常准则。

第三,树立“为进步喝彩”的积极舆论导向。评价改革的深化,离不开外部环境的支持。教育行政部门和社会媒体应协同发力,积极发掘并宣传那些在增值评价实践中取得实质性成效的典型,尤其要关注基础薄弱但进步突出的案例。这有助于在全社会范围内树立“不比基础比进步,不看分数看成长”的健康发展观,进一步推动教育评价从“证明”(PROVE)转向“发展”(IMPROVE) [30]。

## 参考文献

- [1] 陈平静,孙宽宁. 增值评价中国化的历程、羁绊与超越[J]. 国家教育行政学院学报, 2024(11): 79-87.
- [2] 辛涛,周园. 增值评价改革的现状、问题与未来进路[J]. 中小学管理, 2025(10): 9-14.
- [3] 深化新时代教育评价改革总体方案[N]. 人民日报, 2020-10-14(001).
- [4] 中华人民共和国教育部. 义务教育课程方案(2022年版)[M]. 北京: 北京师范大学出版社, 2022: 5.
- [5] 章勇,邹良,刘先发. 新高考增值评价两种模型估计效果的比较研究[J]. 中国考试, 2023(9): 60-69.
- [6] 林迪,彭小敏,周玲.“进步值”激活教育生态: 县域增值评价创新的武侯实践[J]. 中小学管理, 2025(10): 18-21.
- [7] 王家祺,周园,王梦璐. 区域教育质量增值评价结果运用探析——以济南市部分区域学生增值评价为例[J]. 中国考试, 2023(3): 52-59.
- [8] 王光秋. 基于区域四维评价体系的初中绿色增值评价[J]. 教育测量与评价, 2018(3): 34-38.
- [9] 刘邦奇,朱广袤,张金霞. 智能技术支持的增值评价模式及典型实践[J]. 中国远程教育, 2022(12): 49-57.
- [10] 邓澜,赵忠,白鸣珂. 增值评价的校本化设计与整体性实践[J]. 中小学管理, 2025(10): 22-25.
- [11] 郭宝彬.“增值性”综合素质评价: 让学生看见自己的成长[J]. 中小学管理, 2025(10): 26-28.
- [12] 任敏. 作文自能评改: 指向核心素养的增值评价[J]. 语文建设, 2023(20): 32-36.
- [13] 吴国荣. 小学生科学“四项学力”增值评价校本化实践[J]. 教学与管理, 2024(8): 63-67.

- [14] 李锋, 沈玲霞, 林众. 信息科技新课标中素养导向的增值性学习评价[J]. 课程·教材·教法, 2022, 42(11): 44-50.
- [15] Sanders, W.L. and Horn, S.P. (1994) The Tennessee Value-Added Assessment System (TVAAS): Mixed-Model Methodology in Educational Assessment. *Journal of Personnel Evaluation in Education*, **8**, 299-311. <https://doi.org/10.1007/bf00973726>
- [16] Sanders, W.L. and Horn, S.P. (1998) Research Findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for Educational Evaluation and Research. *Journal of Personnel Evaluation in Education*, **12**, 247-256. <https://doi.org/10.1023/a:1008067210518>
- [17] Rothstein, J. (2010) Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. *Quarterly Journal of Economics*, **125**, 175-214. <https://doi.org/10.1162/qjec.2010.125.1.175>
- [18] Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. and Rothstein, J. (2012) Evaluating Teacher Evaluation. *Phi Delta Kappan*, **93**, 8-15. <https://doi.org/10.1177/003172171209300603>
- [19] Morganstein, D. and Wasserstein, R. (2014) ASA Statement on Value-Added Models. *Statistics and Public Policy*, **1**, 108-110. <https://doi.org/10.1080/2330443x.2014.956906>
- [20] American Educational Research Association (2015) AERA Statement on Use of Value-Added Models (VAM) for the Evaluation of Educators and Educator Preparation Programs. *Educational Researcher*, **44**, 448-452.
- [21] 王天平, 牌代琼. 新时代教育增值评价改革: 从数据描绘走向价值呈现[J]. 中国考试, 2022(10): 31-38.
- [22] Chetty, R., Friedman, J.N. and Rockoff, J.E. (2014) Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, **104**, 2633-2679. <https://doi.org/10.1257/aer.104.9.2633>
- [23] 张亮. 美国学校效能增值模型研究的进展与趋势[J]. 教育研究, 2015, 36(11): 135-144.
- [24] 张晋, 马飞. 中小学增值评价的价值向度和实践路径[J]. 课程·教材·教法, 2024, 44(12): 47-54.
- [25] 李俊飞, 谭顶良. 增值评价的基本理念、实践困境与优化策略[J]. 中国教育学刊, 2024(9): 37-43.
- [26] 杨立军, 夏紫微. 教育增值评价 50 年: 演进、挑战与进路[J]. 高教发展与评估, 2024, 40(2): 1-18, 119.
- [27] 陈丽, 林文静, 郑勤华. 人工智能赋能中小学生综合素质评价的创新探索[J]. 中国考试, 2025(2): 41-49.
- [28] 许海霞, 王蕊, 马陆亭. 教育评价改革的几个关键问题[J]. 中国考试, 2020(8): 21.
- [29] 姬国君, 李佳莉, 赵志雯. 综合素质评价“量化”倾向: 实然审思与应然进路[J]. 教育导刊, 2024(10): 63-70.
- [30] 秦春华. 充分发挥教育评价在教育强国建设中的枢纽作用[J]. 华东师范大学学报(教育科学版), 2025, 43(9): 40-53.